

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 13:14:58 ; Search time 56 Seconds
(without alignments)
671.050 Million cell updates/sec

Title: US-09-763-335-2

Perfect score: 754

Sequence: 1 MAMVSAMSWLWLWISACAM.....SGWCAIGNKIKYTRIHPRT 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- A Geneseq_29Jan04.*
- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	754	100.0	133	3	AAY70400 Human cel
2	754	100.0	133	4	AAM78667 Human pro
3	754	100.0	190	4	AAM79651 Human pro
4	746	98.9	133	4	ABUS2838 Human pro
5	432	57.3	131	4	AAM39191 Human pol
6	432	57.3	242	4	AAM40977 Human pol
7	241	32.0	121	3	AAB44958 Human sec
8	241	32.0	132	3	AAY53016 Human sec
9	240.5	31.9	89	4	ABUS2839 Human kid
10	221.5	29.4	115	4	AAM24360 Human EST
11	150	19.9	29	2	AAY13167 Human sec
12	110.5	14.7	1202	7	ADE60833 Rat Prote
13	96	12.7	1964	2	AAM95557 Mus muscu
14	95.5	12.7	2531	7	ADE63713 Rat Prote
15	95.5	12.7	2531	7	ADE63705 Rat Prote
16	95.5	12.7	2531	7	ADE63709 Rat Prote
17	95.5	12.7	2531	7	ADE63701 Rat Prote
18	92.5	12.3	2146	4	ABBE62317 Drosophil
19	91.5	12.1	233	2	AAY29053 T. gondii
20	91.5	12.1	233	4	AAY25524 T. gondii
21	87	11.5	188	4	AAM18112 Novel hum
22	87	11.5	188	4	AAM17021 Human nov
23	87	11.5	188	4	ABBI0465 Human cDN
24	87	11.5	188	4	AAM19912 Novel hum
25	87	11.5	188	5	ABJ05739 Novel hum

ALIGNMENTS

RESULT 1
AAY70400

ID AAY70400 standard; protein; 133 AA.

XX AC AAY70400;

XX DT 21-JUN-2000 (first entry)

XX DE Human cell-signalling protein-2.

XX KW Human cell-signalling protein-2; CSIG-2; neoplastic disorder; neurological disorder; immunological disorder; smooth muscle disorder; vesicle trafficking disorder; asthma; emphysema; rheumatoid arthritis; HIV; human immunodeficiency virus; osteoporosis; multiple sclerosis; arteriosclerosis; diabetes mellitus; psoriasis; viral infection; irritable bowel syndrome; bacterial infection; fungal infection; cytostatic; antiarthritic; antidiabetic; immunosuppressive; antiarteriosclerotic; anti-HIV; antidiabetic; antiinflammatory; neuroprotective; antipsoriatic; antimicrobial; drug screening.

XX OS Homo sapiens.

XX EH Key

XX FT Peptide

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX FT Modified-site

XX WPI; 2000-246562/21.
 DR N-PSDB; AAZ51528.
 XX
 PT New human cell signaling proteins and polynucleotides useful for
 PT diagnosis, prevention and treatment of neoplastic, neurological,
 PT immunological, vesicle trafficking and smooth muscle disorders.
 XX
 PS Claim 1; Page 60; 62pp; English.
 XX
 CC The present sequence is human cell-signalling protein-2 (CSIG-2) involved
 CC in regulation of cell proliferation, differentiation and gene
 CC transcription. CSIG can be used in the diagnosis and treatment of
 CC diseases associated with expression of CSIG. These diseases include
 CC neoplastic, neurological, immunological, vesicle trafficking and smooth
 CC muscle disorders, including HIV, rheumatoid arthritis, asthma,
 CC atherosclerosis, diabetes mellitus, emphysema, irritable bowel syndrome,
 CC multiple sclerosis, osteoporosis, psoriasis and infections including
 CC viral, bacterial and fungal. CSIG and its catalytic or immunogenic
 CC fragments are useful for drug screening using libraries of compounds
 XX
 SQ Sequence 133 AA;
 Query Match 100.0%; Score 754; DB 3; Length 133;
 Best Local Similarity 100.0%; Pred. No. 1.9e-69;
 Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAMVSAMSWLYLWISACAMLLCHGSLQHTFOQHLHRRPGGTCEVIAAHRCCNKNRIE 60
 Db 1 MAMVSAMSWLYLWISACAMLLCHGSLQHTFOQHLHRRPGGTCEVIAAHRCCNKNRIE 60
 QY 61 RSQTVKSCCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCAT 120
 Db 61 RSQTVKSCCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCAT 120
 QY 121 GNKIKTRIHPT 133
 Db 121 GNKIKTRIHPT 133
 RESULT 2
 AAM78667
 ID AAM78667 standard; protein; 133 AA.
 XX
 AC AAM78667;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 1329.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US0004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX
 PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI; 2001-476283/51.
 DR N-PSDB; AAK51800.
 XX
 CC Nucleic acids encoding polypeptides with cytokine-like activities, useful
 CC in diagnosis and gene therapy.
 XX
 PS Claim 20; Page 3574; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3686 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 133 AA;
 Query Match 100.0%; Score 754; DB 4; Length 133;
 Best Local Similarity 100.0%; Pred. No. 1.9e-69;
 Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAMVSAMSWLYLWISACAMLLCHGSLQHTFOQHLHRRPGGTCEVIAAHRCCNKNRIE 60
 Db 1 MAMVSAMSWLYLWISACAMLLCHGSLQHTFOQHLHRRPGGTCEVIAAHRCCNKNRIE 60
 QY 61 RSQTVKSCCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCAT 120
 Db 61 RSQTVKSCCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCAT 120
 QY 121 GNKIKTRIHPT 133
 Db 121 GNKIKTRIHPT 133
 RESULT 3
 AAM79651
 ID AAM79651 standard; protein; 190 AA.
 XX
 AC AAM79651;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 3297.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US0004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.

```
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
DR N-PSDB; AAK52784.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
XX Claim 20; Page 303; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
XX Sequence 190 AA;
SQ
Query Match 100.0%; Score 754; DB 4; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.8e-69;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAMVSAMSWVLYLWISACAMLLCHGSLQHTFOQHLHRRPEGGTCEVIAAHRCCNKNRIE 60
DB 58 MAMVSAMSWVLYLWISACAMLLCHGSLQHTFOQHLHRRPEGGTCEVIAAHRCCNKNRIE 117
QY 61 RSQTVKCSCLPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCAT 120
DB 118 RSQTVKCSCLPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCAT 177
QY 121 GNKIKTRIHPRT 133
DB 178 GNKIKTRIHPRT 190
RESULT 4
ABU52838
ID ABU52838 standard; protein; 133 AA.
XX
XX ABU52838;
XX
XX 14-APR-2003 (first entry)
XX
XX Human kidney-derived protein from DKFZphfkd2_4b6.
XX
XX Human; gene therapy; vaccine; disease treatment; detection.
XX
XX Homo sapiens.
XX
XX WO200112659-A2.
XX
XX 22-FEB-2001.
XX
XX 18-AUG-2000; 2000WO-IB001496.
XX
XX 18-AUG-1999; 99US-0149499P.
PR
```

```
PR 28-SEP-1999; 99US-0156503P.
XX (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
XX Wiemann S;
XX
XX WPI: 2001-327840/34.
DR N-PSDB; ABX71301.
XX
XX Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
XX
XX Claim 21; Page 426; 1095pp; English.
XX
XX This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence represents a
CC polypeptide described in the disclosure of the invention
XX
XX Sequence 133 AA;
SQ
Query Match 98.9%; Score 746; DB 4; Length 133;
Best Local Similarity 99.2%; Pred. No. 1.3e-68;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAMVSAMSWVLYLWISACAMLLCHGSLQHTFOQHLHRRPEGGTCEVIAAHRCCNKNRIE 60
DB 1 MAMVSAMSWVLYLWISACAMLLCHGSLQHTFOQHLHRRPEGGTCEVIAAHRCCNKNRIE 60
QY 61 RSQTVKCSCLPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCAT 120
DB 61 RSQTVKCSCLPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCAT 120
QY 121 GNKIKTRIHPRT 133
DB 121 GNKIKTRIHPRT 133
RESULT 5
AAM39191
ID AAM39191 standard; protein; 131 AA.
XX
XX AAM39191;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 2336.
XX
XX Human; neotrophic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US034263.
XX
XX 23-DEC-1999; 99US-00471275.
XX
XX 21-JAN-2000; 2000US-00488725.
XX
XX 25-APR-2000; 2000US-00552317.
XX
XX 20-JUN-2000; 2000US-00598042.
PR
```

PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AAI56347.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX Example 4; SEQ ID NO 2336; 10078pp; English.
 PS
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 XX Sequence 131 AA;
 Query Match 57.3%; Score 432; DB 4; Length 131;
 Best Local Similarity 61.2%; Pred. No. 2e-36;
 Matches 74; Conservative 14; Mismatches 23; Indels 10; Gaps 1;
 QY 9 WVLYLWISACAMLLCHGSLQHTFQOHLHRRPEGGTCEVIAAHRCCNKNRIERSQTVKCS 68
 Db 19 FIVTLW-----GKVSSANHHKAAHVKTGTCEVVALHRCCKNKNRIERSQTVKCS 68
 QY 69 CLPGKVAGTTRNRPSCVDASIVIGKWCMEBPCLEGECKTLPDNGWMCATGNKIKTR 128
 Db 69 CFPQGVAGTTTAAAPSCVDASIVEQKWCHMQPCLEGECKVLPDRKGWSSGNGKVKTR 128
 QY 129 I 129
 Db 129 V 129
 RESULT 6
 AAM40977
 ID AAM40977 standard; protein; 242 AA.
 AC AAM40977;
 XX
 XX 22-OCT-2001 (first entry)
 DE Human polypeptide SEQ ID NO 5908.
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX Homo sapiens.
 OS

XX WO200153312-A1.
 XX 26-JUL-2001.
 XX 26-DEC-2000; 2000WO-US034263.
 XX 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00553117.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AAI60133.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX Example 2; SEQ ID NO 5908; 10078pp; English.
 PS
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 XX Sequence 242 AA;
 Query Match 57.3%; Score 432; DB 4; Length 242;
 Best Local Similarity 61.2%; Pred. No. 3.9e-36;
 Matches 74; Conservative 14; Mismatches 23; Indels 10; Gaps 1;
 QY 9 WVLYLWISACAMLLCHGSLQHTFQOHLHRRPEGGTCEVIAAHRCCNKNRIERSQTVKCS 68
 Db 130 FIVTLW-----GKVSSANHHKAAHVKTGTCEVVALHRCCKNKNRIERSQTVKCS 179
 QY 69 CLPGKVAGTTRNRPSCVDASIVIGKWCMEBPCLEGECKTLPDNGWMCATGNKIKTR 128
 Db 180 CFPQGVAGTTTAAAPSCVDASIVEQKWCHMQPCLEGECKVLPDRKGWSSGNGKVKTR 239
 QY 129 I 129
 Db 240 V 240
 RESULT 7
 AAB44958
 ID AAB44958 standard; protein; 121 AA.
 AC AAB44958;
 XX
 XX 12-FEB-2001 (first entry)
 DT

QY	6	AMSWVLWLTISACAMLLCHGSLQHTFQQHHLHRPBGCTCEVTAARCCNKNRIEERSQTV	65
	:	: : :	:
Db	6	SMSSTFWAFMILASLIAYCS-----QLAAGTCEIVTLDRDSQPRTTIARQTA	54
	:	: : :	:
QY	66	KCSSLGLGVAGTTNRNPSCVDASIVIGKWCMEPCLEGECKTLPDMSGMWCA-TGNKI	124
	:	: : :	:
Db	55	RCACKRGQIAGTTRAPACVDARIINTKWCDMLPCLEGEGCDLLINRSGWTCTOPGGRI	114
	:	: : :	:
QY	125	KTTRI 129	
	:	: : :	:
Db	115	KTTTV 119	
	:	: : :	:
RESULT 8			
AAVS3016	ID	AAVS3016 standard; protein; 132 AA.	
XX	AC	AAVS3016;	
XX	DT	29-FEB-2000 (first entry)	
XX	DE	Human secreted protein clone ml243_1 protein sequence SEQ ID NO:38.	
XX	KW	Human; secreted protein; nutritional; cytokine; cell proliferation;	
XX	KW	differentiation; immune stimulating; vaccine; suppression;	
KW	KW	haematopoiesis regulation; tissue growth; activin; inhibin; chemotactic;	
KW	KW	chemokinetic; haemostatic; thrombolytic; receptor; ligand;	
KW	KW	anti-inflammation; cadherin; tumour invasion suppressor;	
KW	KW	tumour inhibition; gene therapy.	
XX	OS	Homo sapiens.	
XX	PN	WO9571132-A1.	
XX	XX	11-NOV-1999.	
PD	XX		
PF	07-MAY-1999;	99WO-US009970.	
XX	XX		
PR	07-MAY-1998;	98US-0084564P.	
PR	02-JUN-1998;	98US-0087645P.	
PR	22-JUL-1998;	98US-0093712P.	
PR	31-JUL-1998;	98US-0094935P.	
PR	10-AUG-1998;	98US-0095880P.	
PR	11-AUG-1998;	98US-0096068P.	
PR	06-MAY-1999;	99US-00306111.	
XX	PA	(GENY) GENETICS INST INC.	
XX	PI	Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;	
PI	Merberg D, Treacy M, Agostino MJ, Steininger RJ, Bowman MR;		
PI	Diblasio-Smith E, Widom A;		
XX	XX		
DR	WIPI; 2000-052937/04.		
DR	N-PSDB; AAZ33334.		
XX	XX	New polynucleotides encoding secreted human proteins, derived from adult placenta, adult retina, fetal brain, fetal.	
Claim 47;	Page 397;	492pp; English.	
The present invention describes new human secreted proteins which were isolated from adult placenta, adult retina, foetal brain, foetal kidney, adult blood, adult brain, adult thyroid, adult bladder, adult neural tissue, adult testes, and adult lymph node cDNA libraries. The human secreted proteins, and the polynucleotides encoding them, are predicted to have biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals. Suggested activities include nutritional activity, cytokine and cell proliferation/differentiation activity, immune stimulating (e.g. as vaccines) or suppressing activity, haematopoiesis regulating activity, tissue growth activity, activin/inhibin activity, haemostatic and thrombolytic activity, chemotactic/chemokinetic activity, anti-inflammatory activity, cadherin/tumour receptor/ligand activity, anti-inflammation activity, immunostimulating (e.g. as vaccines) or suppressing activity, haematopoiesis regulating activity,			

CC invasion suppressor activity, and tumour inhibition activity. The
CC polynucleotides are also stated to be useful for gene therapy. AA233316
CC to AA233373 encode human secreted proteins, and AA52998 to AA53060
CC represent human secreted proteins, given in the present invention
XX
SQ Sequence 132 AA;

Query Match 32.0%; Score 241; DB 3; Length 132;
Best Local Similarity 38.4%; Pred. No. 7.9e-17;
Matches 48; Conservative 23; Mismatches 42; Indels 12; Gaps 2;

QY 6 AMSVLYLWISACAMLLCHGSLQHTFOQHLHHPGEGTCVIAAHRCCNKNRIEERSQTV 65
DB 18 SMSSTFWAFMLASLLIAYCS-----QLAAGTCVIVLDRDSSQPRRTIARQTA 66
QY 66 KCSCLPKQVAGTTNRPSQVDASIVIGKWCMEPCLEGECKTLPDNSGWMCA-TGNKI 124
DB 67 RCACRGQIAGTTTRARPACVDARIITKQWCDMLPCLEGECDLLINRSQWCTQPGGRI 126
QY 125 KTTRI 129
DB 127 KTTTV 131

RESULT 9
ID ABUS2839 standard; protein; 89 AA.
AC ABUS2839;

DT 14-APR-2003 (first entry)
DE Human kidney-derived protein DKPZphkd2_4b6 homologue.
KW Human; gene therapy; vaccine; disease treatment; detection.
XX Homo sapiens.
OS
XX WO200112659-A2.
XX 22-FEB-2001.

XX 18-AUG-2000; 2000WO-1B001496.
XX 18-AUG-1999; 99US-0149499P.
XX 28-SEP-1999; 99US-0156503P.
XX (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX Wiemann S;
XX WPI; 2001-327840/34.
XX Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
XX
XX Example III; Page 426; 1095pp; English.

XX This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence represents a homologue
CC of a polypeptide described in the disclosure of the invention
XX
SQ Sequence 89 AA;

Query Match 31.9%; Score 240.5; DB 4; Length 89;
Best Local Similarity 49.4%; Pred. No. 5.7e-17;

Matches 44; Conservative 14; Mismatches 30; Indels 1; Gaps 1;
QY 42 GTCEVIAAHRCCNKNRIEERSQTVKGCCLPGKVAGTTNRPSQVDASIVIGKWCMEPC 101
DB 1 GTCEIVTLDRDSSQPRRTIARQTAACACRGQIAGTTTRARPACVDARIITKQWCDMLPC 60
QY 102 LEGECKTLPDNSGWMCA-TGNKIKTTRI 129
DB 61 LEGEGCDLLINRSQWCTQPGGRIKTTTV 89

RESULT 10
ID AM24360 standard; protein; 115 AA.
XX
AC AM24360;

DT 12-OCT-2001 (first entry)

DE Human EST encoded protein SEQ ID NO: 1885.
XX
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
KW gene therapy; nutrition.
XX Homo sapiens.
OS
XX WO200154477-A2.
XX 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US002687.

XX 25-JAN-2000; 2000US-00491404.
PR 17-JUL-2000; 2000US-00617746.
PR 03-AUG-2000; 2000US-00631451.
PR 15-SEP-2000; 2000US-00663870.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Wehrman T;

XX WPI; 2001-476164/51.
XX N-PSDB; AAH99019.

XX Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use.

XX Claim 20; Page 1217-1218; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit, fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention
XX

SQ Sequence 115 AA;

Query Match 29.4%; Score 221.5; DB 4; Length 115;
Best Local Similarity 54.0%; Pred. No. 6.8e-15;
Matches 47; Conservative 6; Mismatches 21; Indels 13; Gaps 2;

QY 5 SAMSVL-----YLAWSACAMLLCHGSLQHTFOQHLHHPGEGTCVIAAHRCCNKNRIE 59
DB 10 STGGWLLALCLALWLTHTLTLAALQPTATVLVQQ-----GTCEVIAAHRCCNKNRIE 61
QY 60 ERSQTVKGCCLPGKVAGTTNRPSQVD 86

the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene therapy). The sequence presented is a rat protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 2531 AA;

Query Match 12.7%; Score 95.5; DB 7; Length 2531;
Best Local Similarity 26.8%; Pred. No. 1.7;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;
QY 15 ISACAMLLCHGSLQHTFQQHHLRPEGTCR-VIAHRC-----CNKN 56
Db 679 IDECAGSPCH-----NGGTCEGDIAGFTCRCPGTYHPTCLSEVNECNSN 723
QY 57 -----RIERSQTVKCSCLPGKVGATRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
Db 724 PCIHGACRDGLNGYKDCAPG-WSGTNCD-----INNNECESNPCVNGGTCKDM- 771
QY 112 DNSGMMC 118
Db 772 -TSGYVC 777

RESULT 15

AD563705
ID ADE63705 standard; protein; 2531 AA.
AC ADE63705;
XX 29-JAN-2004 (first entry)
DT Rat Protein CAA40667, SEQ ID NO 9649.
DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX Rattus norvegicus.
OS WO2003016475-A2.
XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WEI; 2003-268312/26.
DR GENBANK; CAA40667.
XX New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

Claim 1; Page; 1017pp; English.
PS The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene therapy). The sequence presented is a rat protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 2531 AA;

Query Match 12.7%; Score 95.5; DB 7; Length 2531;
Best Local Similarity 26.8%; Pred. No. 1.7;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;
QY 15 ISACAMLLCHGSLQHTFQQHHLRPEGTCR-VIAHRC-----CNKN 56
Db 679 IDECAGSPCH-----NGGTCEGDIAGFTCRCPGTYHPTCLSEVNECNSN 723
QY 57 -----RIERSQTVKCSCLPGKVGATRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
Db 724 PCIHGACRDGLNGYKDCAPG-WSGTNCD-----INNNECESNPCVNGGTCKDM- 771
QY 112 DNSGMMC 118
Db 772 -TSGYVC 777

Search completed: April 8, 2004, 13:16:11
Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 13:14:59 ; Search time 18 Seconds
(without alignments)
384.741 Million cell updates/sec

Title: US-09-763-335-2

Perfect score: 754

Sequence: 1 MAMVMSWVLYLWISACAM.....SCHWMCATGNKIXKTRHPRT 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110.5	14.7	1202	JAG2 RAT	P97607 rattus norv
2	96	12.7	1964	NTC4 MOUSE	P31695 mus musculu
3	95.5	12.7	2531	NTC1 MOUSE	Q01705 mus musculu
4	95.5	12.7	2531	NTC1 RAT	Q07008 rattus norv
5	92.5	12.3	2139	CRB DROME	P10040 drosophila
6	91.5	12.1	1213	JAG3 BRARE	Q90Y54 brachydanio
7	91	12.1	555	DP87 DICDI	Q04503 dictyosteli
8	86.5	11.5	2703	NOTC DROME	P07207 drosophila
9	86	11.4	685	DLA4 HUMAN	Q9N161 homo sapien
10	86	11.4	1218	JAG1 HUMAN	P78504 homo sapien
11	86	11.4	1218	JAG1 MOUSE	Q9QXX0 mus musculu
12	86	11.4	1219	JAG1 RAT	Q63722 rattus norv
13	85.5	11.3	1064	FBP1 STRPU	P10079 strongyloce
14	84.5	11.2	2003	NTC4 HUMAN	Q99466 homo sapien
15	84	11.1	1247	JAG2 MOUSE	Q9GYE5 mus musculu
16	84	11.1	2524	NOTC XENLA	P21783 xenopus lae
17	83.5	11.1	1170	TSP2 BOVIN	Q95116 bos taurus
18	82.5	10.9	2470	NTC2 MOUSE	Q35516 mus musculu
19	82.5	10.9	2471	NTC2 HUMAN	Q04721 homo sapien
20	82.5	10.9	2471	NTC2 RAT	Q9QW30 rattus norv
21	82	10.9	2318	NTC3 MOUSE	Q61982 mus musculu
22	82	10.9	4544	LRP1 HUMAN	Q07954 homo sapien
23	82	10.9	5376	ZAN MOUSE	Q88799 mus musculu
24	81.5	10.8	355	YMT5 YEAST	Q04018 saccharomyc
25	81.5	10.8	2556	NTC1 HUMAN	P46531 homo sapien
26	80.5	10.7	2321	NTC3 HUMAN	Q9UM47 homo sapien
27	80	10.6	937	VWF BOVIN	P80012 bos taurus
28	79.5	10.5	2319	NTC3 RAT	Q9R172 rattus norv
29	79	10.5	819	AD09 HUMAN	Q13443 homo sapien
30	78.5	10.4	1178	TSP2 CHICK	P35440 gallus gall
31	78.5	10.4	3312	CLK3 HUMAN	Q9NVQ7 homo sapien
32	78	10.3	1191	LMG2 MOUSE	Q61092 mus musculu
33	78	10.3	1376	CRBH HUMAN	P82279 homo sapien

34	78	10.3	2311	1	FBM2 HUMAN	P35556 homo sapien
35	78	10.3	4543	1	LRP1 CHICK	P98157 gallus gall
36	77.5	10.3	686	1	DLA4 MOUSE	Q9J171 mus musculu
37	77.5	10.3	723	1	DLA1 HUMAN	Q00548 homo sapien
38	77.5	10.3	1125	1	TIR2 BOVIN	Q06807 bos taurus
39	77.5	10.3	1877	1	PKX5 MOUSE	Q04592 mus musculu
40	77.5	10.3	2437	1	NTC1 BRARE	P46530 brachydanio
41	77	10.2	843	1	CO7 HUMAN	P10643 homo sapien
42	77	10.2	1403	1	NID2 MOUSE	Q88322 mus musculu
43	77	10.2	2907	1	FBM2 MOUSE	Q61555 mus musculu
44	76.5	10.1	1172	1	TSP2 MOUSE	Q03350 mus musculu
45	76	10.1	592	1	DLA3 MOUSE	Q88516 mus musculu

ALIGNMENTS

RESULT 1				
JAG2 RAT	STANDARD;	PRT;	1202 AA.	
ID	JAG2 RAT			
AC	P97607;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Jagged 2 (Jagged2) (fragment).			
GN	JAG2.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=97105852; PubMed=8948600;			
RA	Shawber C., Boulter J., Lindsell C.E., Weinmaster G.;			
RT	"Jagged2: a serrate-like gene expressed during rat embryogenesis.";			
RL	Dev. Biol. 180:370-376(1996).			
CC	FUNCTION: Putative Notch ligand involved in the mediation of Notch signaling. May have a role in neurogenesis in the peripheral nervous system, limb development and in the adult brain.			
CC	SUBCELLULAR LOCATION: Type I membrane protein.			
CC	DEVELOPMENTAL STAGE: At stage E12.5 it is detected in dorsal root ganglia, AER, and surface ectoderm. At E14.5, found as well in cranial ganglia, thymus and olfactory epithelia. At E16.5, found as well in salivary gland, tooth buds and hair follicles.			
CC	SIMILARITY: Contains 16 EGF-like domains.			
CC	SIMILARITY: Contains 1 DSL domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; U70050; AAC52946.1; -.			
DR	HSSP; P00743; 1CCF.			
DR	GO; GO:0005887; C:integral to plasma membrane; ISS.			
DR	GO; GO:0008083; P:growth factor activity; ISS.			
DR	GO; GO:0005112; P:Notch binding; NAS.			
DR	GO; GO:0007049; P:cell cycle; ISS.			
DR	GO; GO:0001054; P:cell differentiation; NAS.			
DR	GO; GO:0001709; P:cell fate determination; NAS.			
DR	GO; GO:0009912; P:hair cell fate commitment; ISS.			
DR	GO; GO:0007605; P:hearing; ISS.			
DR	GO; GO:0030326; P:limb morphogenesis; NAS.			
DR	GO; GO:0007219; P:N signaling pathway; NAS.			
DR	GO; GO:0030334; P:regulation of cell migration; ISS.			
DR	GO; GO:0042127; P:regulation of cell proliferation; ISS.			
DR	GO; GO:0007283; P:spermatogenesis; IEP.			
DR	GO; GO:0030217; P:T-cell differentiation; ISS.			
DR	GO; GO:0045061; P:thymic T-cell selection; ISS.			

DR InterPro; IPR000152; Asx_hydroxyl_8.
 DR InterPro; IPR000174; DSL.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001438; EGF_II.
 DR InterPro; IPR006209; EGF_Like.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF01414; DSL; 1.
 DR Pfam; PF00008; EGF; 14.
 DR PRINTS; PR00010; EGFBLD.
 DR SMART; SM00051; DSL; 1.
 DR SMART; SM00179; EGF_CA; 7.
 DR SMART; SM00214; VMC7; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 10.
 DR PROSITE; PS00022; EGF_1; 15.
 DR PROSITE; PS01186; EGF_2; 11.
 DR PROSITE; PS00026; EGF_3; 15.
 DR PROSITE; PS01187; EGF_CA; 7.
 KW Calcium-binding; EGF-Like domain; Glycoprotein; Developmental protein;
 KW Repeat; Transmembrane.
 FT NON_TER 1 1
 FT DOMAIN <1 1085
 FT TRANSMEM 1086 1102
 FT DOMAIN 1103 1202
 FT DOMAIN 132 194
 FT DOMAIN 195 228
 FT DOMAIN 229 259
 FT DOMAIN 261 299
 FT DOMAIN 301 337
 FT DOMAIN 339 375
 FT DOMAIN 377 413
 FT DOMAIN 415 450
 FT DOMAIN 452 488
 FT DOMAIN 490 527
 FT DOMAIN 529 589
 FT DOMAIN 591 627
 FT DOMAIN 629 665
 FT DOMAIN 667 703
 FT DOMAIN 706 742
 FT DOMAIN 744 780
 FT DOMAIN 782 818
 FT DISULFID 199 210
 FT DISULFID 203 216
 FT DISULFID 218 227
 FT DISULFID 230 241
 FT DISULFID 236 247
 FT DISULFID 249 258
 FT DISULFID 265 277
 FT DISULFID 271 287
 FT DISULFID 289 298
 FT DISULFID 305 316
 FT DISULFID 310 325
 FT DISULFID 327 336
 FT DISULFID 343 354
 FT DISULFID 348 363
 FT DISULFID 365 374
 FT DISULFID 381 392
 FT DISULFID 386 401
 FT DISULFID 403 412
 FT DISULFID 419 429
 FT DISULFID 423 438
 FT DISULFID 440 449
 FT DISULFID 456 467
 FT DISULFID 461 476
 FT DISULFID 478 487
 FT DISULFID 495 506
 FT DISULFID 500 515
 FT DISULFID 517 526
 FT DISULFID 544 567
 FT DISULFID 561 577
 FT DISULFID 579 588
 FT DISULFID 595 606
 FT DISULFID 600 615

FT DISULFID 617 626 BY SIMILARITY.
 FT DISULFID 633 644 BY SIMILARITY.
 FT DISULFID 638 653 BY SIMILARITY.
 FT DISULFID 655 664 BY SIMILARITY.
 FT DISULFID 671 682 BY SIMILARITY.
 FT DISULFID 676 691 BY SIMILARITY.
 FT DISULFID 693 702 BY SIMILARITY.
 FT DISULFID 710 721 BY SIMILARITY.
 FT DISULFID 715 730 BY SIMILARITY.
 FT DISULFID 732 741 BY SIMILARITY.
 FT DISULFID 748 759 BY SIMILARITY.
 FT DISULFID 753 768 BY SIMILARITY.
 FT DISULFID 770 779 BY SIMILARITY.
 FT DISULFID 786 797 BY SIMILARITY.
 FT DISULFID 791 806 BY SIMILARITY.
 FT DISULFID 808 817 BY SIMILARITY.
 SQ SEQUENCE 1202 AA; 129703 MW; 08CB4E5271FF8BE CRC64;
 Query Match 14.7%; Score 110.5; DB 1; Length 1202;
 Best Local Similarity 27.0%; Pred. No. 0.0029;
 Matches 33; Conservative 13; Mismatches 27; Indels 49; Gaps 9;
 QY 23 CHGSLQHTFOQHHLRPEGGTC-EVIAAHC-----CNKNRI- 58
 DB 419 CHGQCH-----GCTCKDLVNGVQVCYPRGFGGRHCELEYKCASSPCRRGGIC 467
 QY 59 ERSQTVKSCSLPGKVGAGTTRNRPS-CVDASIVIGKWCMEPCLEGEBECKTLPNDSGM 117
 DB 468 EDLVGDFRCHC-PRGLSG-----PLCEVDVLDL-----WCEPNPCLNGARCYNLEDD--YY 514
 QY 118 CA 119
 DB 515 CA 516
 RESULT 2
 NTC4 MOUSE
 ID NTC4 MOUSE STANDARD; PRT; 1964 AA.
 AC P31695; O35442; O88314; O88316; Q62390; Q9R1W9; Q9R1X0;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neurogenic locus notch homolog protein 4 precursor (Notch 4)
 DE [Contains: Transforming protein Int-3].
 GN NOTCH4 OR INT3 OR INT-3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92194507; PubMed=1312643;
 RA Robbins J., Blondel B.J., Gallahan D., Callahan R.;
 RT "Mouse mammary tumor gene int-3: a member of the notch gene family
 RT transforms mammary epithelial cells.";
 RL J. Virol. 66:2594-2599(1992).
 RN [2]
 RP REVISIONS, SEQUENCE FROM N.A.
 RX MEDLINE=97294599; PubMed=9150355;
 RA Gallahan D., Callahan R.;
 RT "The mouse mammary tumor associated gene INT3 is a unique member of
 RT the NOTCH gene family (NOTCH4).";
 RL Oncogene 14:1893-1896(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung, and Testis;
 RX MEDLINE=96281668; PubMed=8681805;
 RA Uyttendaele H., Marazzi G., Wu G., Yan Q., Sassoon D., Kitajewski J.;
 RT "Notch4/int-3, a mammary proto-oncogene, is an endothelial
 RT cell-specific mammalian Notch gene.";
 RL Development 122:2251-2259(1996).
 RN [4]
 RP SEQUENCE FROM N.A.

RA Rowen L., Mahairas G., Qin S., Ahearn M.E., Dankers C., Lasky S.,
 RA Loretz C., Schmidt S., Tipton S., Traicoff R., Zackrone K., Hood L.,
 RT "Sequence of the mouse major histocompatibility locus class III
 RL region.";
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 1436-1600 FROM N.A.
 RX MEDLINE=99252212; PubMed=10233982;
 RA Lee J.-S., Haruna T., Ishimoto A., Honjo T., Yanagawa S.-I.,
 RT "Intracisternal type A particle-mediated activation of the Notch4/int3
 RT gene in a mouse mammary tumor: generation of truncated Notch4/int3
 RT mRNAs by retroviral splicing events.";
 RL J. Virol. 73:5166-5171(1999).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=21244657; PubMed=11344305;
 RA Uyttendaele H., Ho J., Rossant J., Kitajewski J.,
 RT "Vascular patterning defects associated with expression of activated
 RT Notch4 in embryonic endoderm.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:5643-5648(2001).
 RN [7]
 RP SEQUENCE OF 1463-1964, POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS
 OF VAL-1463.
 RX MEDLINE=215233956; PubMed=11518718;
 RA Saxena M.I., Schroeter E.H., Mumm J.S., Kopan R.,
 RT "Murine notch homologs (N1-4) undergo presenilin-dependent
 RT proteolysis.";
 RL J. Biol. Chem. 276:40268-40273(2001).
 RN [8]
 RP POST-TRANSLATIONAL PROCESSING.
 RX MEDLINE=21374376; PubMed=11459941;
 RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.,
 RT "Conservation of the biochemical mechanisms of signal transduction
 RT among mammalian Notch family members.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031(2001).
 CC -!- FUNCTION: Functions as a receptor for membrane-bound ligands
 CC Jagged1, Jagged2 and Deltal to regulate cell-fate determination.
 CC Upon ligand activation through the released notch intracellular
 CC domain (NICD) it forms a transcriptional activator complex with
 CC RBP-J kappa and activates genes of the enhancer of split locus.
 CC Affects the implementation of differentiation, proliferation and
 CC apoptotic programs (By similarity). May regulate branching
 CC morphogenesis in the developing vascular system.
 CC -!- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
 CC terminal fragment N(EC) which are probably linked by disulfide
 CC bonds.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Following
 CC proteolytical processing NICD is translocated to the nucleus.
 CC -!- TISSUE SPECIFICITY: Highly expressed in lung, moderately in heart
 CC kidney, and at lower levels in the ovary and skeletal muscle. A
 CC very low expression is seen in the brain, intestine, liver and
 CC testis.
 CC -!- DEVELOPMENTAL STAGE: Highly expressed in endothelial cells during
 CC embryonic development from 9.0 dpc.
 CC -!- PTM: Synthesized in the endoplasmic reticulum as an inactive form
 CC which is proteolytically cleaved by a furin-like convertase in the
 CC trans-Golgi network before it reaches the plasma membrane to yield
 CC an active, ligand-accessible form. Cleavage results in a C-
 CC terminal fragment N(TM) and a N-terminal fragment N(EC). Following
 CC ligand binding, it is cleaved by TNF-alpha converting enzyme
 CC (TACE) to yield a membrane-associated intermediate fragment called
 CC notch extracellular truncation (NEXT). This fragment is then
 CC cleaved by presenilin dependent gamma-secretase to release a
 CC notch-derived peptide containing the intracellular domain (NICD)
 CC from the membrane.
 CC -!- PTM: Phosphorylated.
 CC -!- DISEASE: Loss of the extracellular domain causes constitutive
 CC activation of the Notch protein, which leads to hyperproliferation
 CC of glandular epithelial tissues and development of mammary
 CC carcinomas.
 CC -!- SIMILARITY: Belongs to the NOTCH family.
 CC -!- SIMILARITY: Contains 29 EGF-like domains.
 CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.

CC -!- SIMILARITY: Contains 5 ANK repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M80456; AAB38377.1; -;
 CC EMBL: U43691; AAC52630.1; -;
 CC EMBL: U43691; AAC52631.1; -;
 CC EMBL: AF030001; AAB82004.1; -;
 CC EMBL: AB016771; BAA32281.1; ALT SEQ.
 CC EMBL: AB016772; BAA32283.1; ALT INIT.
 CC EMBL: AB016773; BAA32284.1; ALT_INIT.
 CC EMBL: AB016774; BAA32285.1; -;
 CC FIR: A38072; TWMV73.
 CC FIR: T09059; T09059.
 CC HSSP: P08709; IBF9.
 CC MGD: MGI:107471; Notch4.
 CC InterPro: IPR002110; ANK.
 CC InterPro: IPR000152; Asx_hydroxyl_S.
 CC InterPro: IPR000742; EGF_2.
 CC InterPro: IPR001881; EGF_Ca.
 CC InterPro: IPR001438; EGF_II.
 CC InterPro: IPR006209; EGF_like.
 CC InterPro: IPR002049; Laminin_EGF.
 CC InterPro: IPR008297; Notch.
 CC InterPro: IPR000800; Notch_dom.
 CC Pfam: PF00023; ank; 6.
 CC Pfam: PF00008; EGF; 27.
 CC Pfam: PF00066; notch; 2.
 CC PIRSF: PIRSF002279; Notch; 1.
 CC PRINTS: PR00010; EGFBL00D.
 CC PRINTS: PR00011; EGFFLAMININ.
 CC PRINTS: PR01452; NOTCH.
 CC SMART: SM00248; ANK; 6.
 CC SMART: SM00179; EGF_CA; 11.
 CC SMART: SM00004; NL; 2.
 CC PROSITE: PS00297; ANK_REPEAT; 1.
 CC PROSITE: PS00088; ANK_REPEAT; 5.
 CC PROSITE: PS00010; ASX_HYDROXYL; 11.
 CC PROSITE: PS00022; EGF_1; 28.
 CC PROSITE: PS01186; EGF_2; 21.
 CC PROSITE: PS00026; EGF_3; 27.
 CC PROSITE: PS01187; EGF_CA; 9.
 CC Receptor; transcription regulation; Activator; Differentiation;
 CC Developmental protein; Repeat; ANK repeat; EGF-like domain;
 CC Transmembrane; Glycoprotein; Signal; Phosphorylation; Proto-oncogene.
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN

FT DOMAIN 586 622 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 623 656 EGF-LIKE 16.
 FT DOMAIN 658 686 EGF-LIKE 17.
 FT DOMAIN 688 724 EGF-LIKE 18.
 FT DOMAIN 726 762 EGF-LIKE 19.
 FT DOMAIN 764 800 EGF-LIKE 20.
 FT DOMAIN 803 839 EGF-LIKE 21.
 FT DOMAIN 841 877 EGF-LIKE 22.
 FT DOMAIN 878 924 EGF-LIKE 23.
 FT DOMAIN 926 962 EGF-LIKE 24.
 FT DOMAIN 964 1000 EGF-LIKE 25.
 FT DOMAIN 1002 1040 EGF-LIKE 26.
 FT DOMAIN 1042 1081 EGF-LIKE 27.
 FT DOMAIN 1083 1122 EGF-LIKE 28.
 FT DOMAIN 1126 1167 EGF-LIKE 29.
 FT REPEAT 1168 1208 LIN/NOTCH 1.
 FT REPEAT 1209 1242 LIN/NOTCH 2.
 FT REPEAT 1243 1282 LIN/NOTCH 3.

Query Match 12.7%; Score 96; DB 1; Length 1964;
 Best Local Similarity 26.8%; Pred. No. 0.13;
 Matches 33; Conservative 8; Mismatches 48; Indels 34; Gaps 6;

QY 15 ISACAMLLC-HGSLQHTFOHHLRPEGGTC-----EVIAAHR--CCNKRI 58
 689 LGGCTSTPCAAGGTCHP-----QPSGYNCPCPAGYGLTCTSEVTACHSGPCLNGGSC 741

QY 59 ERSQTVKSCSLPGKVGATRRNRPCSDVASIVGWKCMEPCLEGECKTLPDNGMWC 118
 742 STRPEGYSCCLP---SHTGRHCQTAVD-----HCVSASCLNGGTCVKNKPGTFFCLC 790

QY 119 ATG 121
 791 ATG 793

RESULT 3

ID NTCL_MOUSE

AC Q01705; Q06007; Q61905; Q990C2; Q90W58; Q90RX7;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenic locus notch homolog protein 1 precursor (Notch 1) (MOTCH A) (wt14) (p300).
 GN NOTCH1 OR MOTCH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Embryo;
 RX MEDLINE=93194170; PubMed=8449489;
 RA Franco del Amo F., Gendron-Maguire M., Swiatek P.J., Jenkins N.A.,
 RA Copeland N.G., Gridley T.;
 RT "Cloning, analysis, and chromosomal localization of Notch-1, a mouse
 RT homolog of Drosophila Notch.";
 RL Homologs 15:259-264(1993).
 RN [2]
 RP SEQUENCE OF 731-1899 FROM N.A. (ISOFORM 2), AND DEVELOPMENTAL STAGE.
 RC STRAIN=CD-1; TISSUE=Embryo;
 RX MEDLINE=93050801; PubMed=1426644;
 RA Raune A.G., Conlon R.A., Ziringibl R., Yamaguchi T.P., Rossant J.;
 RT "Expression analysis of a Notch homologue in the mouse embryo.";
 RL Dev. Biol. 154:377-387(1992).
 RN [3]
 RP SEQUENCE OF 1551-1647 FROM N.A. (ISOFORM 1), AND DEVELOPMENTAL STAGE.
 RC TISSUE=Embryo;
 RX MEDLINE=93048835; PubMed=1425352;
 RA Franco del Amo F., Smith D.E., Swiatek P.J., Gendron-Maguire M.,
 RA Greenspan R.J., McMahon A.P., Gridley T.;
 RT "Expression pattern of Notch, a mouse homolog of Drosophila Notch,

suggests an important role in early postimplantation mouse
 development.";
 RL Development 115:737-744 (1992).
 RN [4]
 RP SEQUENCE OF 1161-1547 FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Embryo;
 RX MEDLINE=93178563; PubMed=8440332;
 RA Lardelli M., Lendahl U.;
 RT "Notch A and Notch B-two mouse Notch homologues coexpressed in a
 RT wide variety of tissues.";
 RL Exp. Cell Res. 204:364-372(1993).
 RN [5]
 RP SEQUENCE OF 1659-1673 FROM N.A.
 RX MEDLINE=99364499; PubMed=10437788;
 RA Lee J.S., Ishimoto A., Yanagawa S.I.;
 RT "Murine leukemia provirus-activated activation of the Notch1 gene leads
 RT to induction of HES-1 in a mouse T lymphoma cell line, DL-3.";
 RL FEBS Lett. 455:276-280(1999).
 RN [6]
 RP SEQUENCE OF 1950-2201 FROM N.A.
 RX MEDLINE=98029496; PubMed=9384671;
 RA Messerle M., Folio M., Nehls M., Eggert H., Boehm T.;
 RT "Dynamic changes in gene expression during in vitro differentiation of
 RT mouse embryonic stem cells.";
 RL Cytokines Cell. Mol. Ther. 1:139-143(1995).
 RN [7]
 RP SEQUENCE OF 1655-1659, CLEAVAGE BY FURIN-LIKE CONVERTASE, AND
 RT MUTAGENESIS OF 1651-ARG--ARG-1654.
 RX MEDLINE=98318619; PubMed=9653148;
 RA Logeat F., Bessia C., Brou C., Lebaill O., Jarriault S., Seidah N.G.,
 RA Israel A.;
 RT "The Notch1 receptor is cleaved constitutively by a furin-like
 RT convertase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8108-8112(1998).
 RN [8]
 RP PARTIAL SEQUENCE, AND POST-TRANSLATIONAL PROCESSING.
 RX MEDLINE=21523956; PubMed=11519718;
 RA Saxena M.T., Schroeter E.H., Mumm J.S., Kopan R.;
 RT "Murine notch homologs (N1-4) undergo presenilin-dependent
 RT proteolysis.";
 RL J. Biol. Chem. 276:40268-40273 (2001).
 RN [9]
 RP POST-TRANSLATIONAL PROCESSING.
 RX MEDLINE=21374376; PubMed=11459941;
 RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.;
 RT "Conservation of the biochemical mechanisms of signal transduction
 RT among mammalian Notch family members.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031(2001).
 RN [10]
 RP INTERACTION WITH DTX1 AND DTX2.
 RX MEDLINE=21123790; PubMed=11226752;
 RA Kishi N., Tang Z., Maeda Y., Hirai A., Mo R., Ito M., Suzuki S.,
 RA Nakao K., Kinoshita T., Kadesch T., Hui C.-C., Artavanis-Tsakonas S.,
 RA Okano H., Matsuno K.;
 RT "Murine homologs of deltex define a novel gene family involved in
 RT vertebrate Notch signaling and neurogenesis.";
 RL Int. J. Dev. Neurosci. 19:21-35(2001).
 CC -!- FUNCTION: Functions as a receptor for membrane-bound ligands
 CC Jagged1, Jagged2 and Deltal to regulate cell-fate determination.
 CC Upon ligand activation through the released notch intracellular
 CC domain (NTCD) it forms a transcriptional activator complex with
 CC RBP-J kappa and activates genes of the enhancer of split locus.
 CC Affects the implementation of differentiation, proliferation and
 CC apoptotic programs (By similarity). May play an essential role in
 CC postimplantation development, probably in some aspect of cell
 CC specification and/or differentiation. May be involved in mesoderm
 CC development, somite formation and neurogenesis. Involved in the
 CC maturation of both CD4+ and CD8+ cells in the thymus.
 CC -!- SUBUNIT: Heterodimer of a C-terminal fragment N(TW) and a N-
 CC terminal fragment N(FC) which are probably linked by disulfide
 CC bonds. Interacts with DTX1 and DTX2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Following
 CC proteolytical processing NICD is translocated to the nucleus.

-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=Q01705-1; Sequences-Displayed;
Name=2;
IsoId=Q01705-2; Sequences-VSP 001402, VSP_001403, VSP_001404;
Note-No experimental confirmation available;
-!- TISSUE SPECIFICITY: Highly expressed in the brain, lung and thymus. Expressed at lower levels in the spleen, bone-marrow, spinal cord, eyes, mammary gland, liver, intestine, skeletal muscle, kidney and heart.
-!- DEVELOPMENTAL STAGE: First detected in the mesoderm at 7.5 dpc By 8.5 dpc highly expressed in presomitic mesoderm, mesenchyme and endothelial cells, while much lower levels are seen in the neuroepithelium. Between 9.5-10.5 dpc expressed at high levels in ectoderm, eye and developing whisker follicles.
-!- PTM: Synthesized in the endoplasmic reticulum as an inactive form which is proteolytically cleaved by a furin-like convertase in the trans-Golgi network before it reaches the plasma membrane to yield an active ligand-accessible form. Cleavage results in a C-terminal fragment N(TM) and a N-terminal fragment N(EC). Following ligand binding, it is cleaved by TNF-alpha converting enzyme (TACE) to yield a membrane-associated intermediate fragment called notch extracellular truncation (NEXT). This fragment is then cleaved by presenilin dependent gamma-secretase to release a notch-derived peptide containing the intracellular domain (NICD) from the membrane.
-!- PTM: Phosphorylated.
-!- SIMILARITY: Belongs to the NOTCH family.
-!- SIMILARITY: Contains 36 EGF-like domains.
-!- SIMILARITY: Contains 3 lin/Notch repeats.
-!- SIMILARITY: Contains 5 ANK repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; Z11886; CAA77941.1; .
EMBL; L02613; AAK14898.1; .
EMBL; X68278; CAA48339.1; .
EMBL; AJ238029; CAB40733.1; .
EMBL; X82562; CAA57909.1; .
PIR; A46019; A46019.
PIR; B49175; B49175.
HSP; P00740; IEDM.
MGD; MG1:97363; Notchl.
GO; GO:0005887; C:integral to plasma membrane; IC.
GO; GO:0005515; F:protein binding; IPI.
GO; GO:0030154; P:cell differentiation; IMP.
GO; GO:0007386; P:compartment specification; IMP.
GO; GO:0007219; P:N signaling pathway; IC.
GO; GO:0045944; P:positive regulation of transcription from P. . ; IDA.
DR InterPro; IPR001152; Asx_hydroxyl_S.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001439; EGF-II.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002049; laminin_EGF.
DR InterPro; IPR008297; Notch.
DR InterPro; IPR000800; Notch_dom.
Pfam; PF00023; ank; 7.
Pfam; PF00008; EGF; 35.
Pfam; PF00066; notch; 3.
PRINTS; PIRSF002279; Notch; 1.
PRINTS; PR00010; EGFELOO.
PRINTS; PR00011; EGFAMIN.
PRINTS; PR01452; NOTCH.

-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=Q01705-1; Sequences-Displayed;
Name=2;
IsoId=Q01705-2; Sequences-VSP 001402, VSP_001403, VSP_001404;
Note-No experimental confirmation available;
-!- TISSUE SPECIFICITY: Highly expressed in the brain, lung and thymus. Expressed at lower levels in the spleen, bone-marrow, spinal cord, eyes, mammary gland, liver, intestine, skeletal muscle, kidney and heart.
-!- DEVELOPMENTAL STAGE: First detected in the mesoderm at 7.5 dpc By 8.5 dpc highly expressed in presomitic mesoderm, mesenchyme and endothelial cells, while much lower levels are seen in the neuroepithelium. Between 9.5-10.5 dpc expressed at high levels in ectoderm, eye and developing whisker follicles.
-!- PTM: Synthesized in the endoplasmic reticulum as an inactive form which is proteolytically cleaved by a furin-like convertase in the trans-Golgi network before it reaches the plasma membrane to yield an active ligand-accessible form. Cleavage results in a C-terminal fragment N(TM) and a N-terminal fragment N(EC). Following ligand binding, it is cleaved by TNF-alpha converting enzyme (TACE) to yield a membrane-associated intermediate fragment called notch extracellular truncation (NEXT). This fragment is then cleaved by presenilin dependent gamma-secretase to release a notch-derived peptide containing the intracellular domain (NICD) from the membrane.
-!- PTM: Phosphorylated.
-!- SIMILARITY: Belongs to the NOTCH family.
-!- SIMILARITY: Contains 36 EGF-like domains.
-!- SIMILARITY: Contains 3 Lin/Notch repeats.
-!- SIMILARITY: Contains 5 ANK repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; Z11886; CAA77941.1; .
EMBL; L02613; AAK14898.1; .
EMBL; X68278; CAA48339.1; .
EMBL; AJ238029; CAB40733.1; .
EMBL; X82562; CAA57909.1; .
PIR; A46019; A46019.
PIR; B49175; B49175.
HSP; P00740; IEDM.
MGD; MG1:97363; Notchl.
GO; GO:0005887; C:integral to plasma membrane; IC.
GO; GO:0005515; F:protein binding; IPI.
GO; GO:0030154; P:cell differentiation; IMP.
GO; GO:0007386; P:compartment specification; IMP.
GO; GO:0007219; P:N signaling pathway; IC.
GO; GO:0045944; P:positive regulation of transcription from P. . ; IDA.
DR InterPro; IPR001152; Asx_hydroxyl_S.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001439; EGF-II.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002049; laminin_EGF.
DR InterPro; IPR008297; Notch.
DR InterPro; IPR000800; Notch_dom.
Pfam; PF00023; ank; 7.
Pfam; PF00008; EGF; 35.
Pfam; PF00066; notch; 3.
PRINTS; PIRSF002279; Notch; 1.
PRINTS; PR00010; EGFELOO.
PRINTS; PR00011; EGFAMIN.
PRINTS; PR01452; NOTCH.

RA Weinmaster G., Roberts V.J., Lenke G.;
 RT "Notch2: a second mammalian Notch gene."
 RN Development 116:931-941(1992).
 RN [5]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=21331789; PubMed=11438922;
 RA Irvin D.K., Zuercher S.D., Nguyen T., Weinmaster G., Kornblum H.I.;
 RT Expression patterns of Notch1, Notch2, and Notch3 suggest multiple
 RT functional roles for the Notch-BSL signaling system during brain
 RT development."
 RT J. Comp. Neurol. 436:167-181(2001).
 CC -!- FUNCTION: Functions as a receptor for membrane-bound ligands
 CC Jagged1, Jagged2 and Delta1 to regulate cell-fate determination.
 CC Upon ligand activation through the released cell-fate determination
 CC domain (NICD) it forms a transcriptional activator complex with
 CC RBP-J kappa and activates genes of the enhancer of split locus.
 CC Affects the implementation of differentiation, proliferation and
 CC apoptotic programs (By similarity). Acts instructively to control
 CC the cell fate determination of CNS multipotent progenitor cells,
 CC resulting in astroglial induction and neuron/oligodendrocyte
 CC suppression.
 CC -!- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
 CC terminal fragment N(EC) which are probably linked by disulfide
 CC bonds (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Following
 CC proteolytical processing NICD is translocated to the nucleus (By
 CC similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in the brain, kidney and spleen.
 CC Expressed in postnatal central nervous system (CNS) germinal zones
 CC and, in early postnatal life, within numerous cells throughout the
 CC CNS. Found in both subventricular and ventricular germinal zones.
 CC -!- DEVELOPMENTAL STAGE: In the embryo, highest levels occur between
 CC days 12 and 14 and decrease rapidly to much lower levels in the
 CC adult.
 CC -!- PTM: Synthesized in the endoplasmic reticulum as an inactive form
 CC which is proteolytically cleaved by a furin-like convertase in the
 CC trans-Golgi network before it reaches the plasma membrane to yield
 CC an active, ligand-accessible form. Cleavage results in a C-
 CC terminal fragment N(TM) and a N-terminal fragment N(EC). Following
 CC ligand binding, it is cleaved by TNF-alpha converting enzyme
 CC (TACE) to yield a membrane-associated intermediate fragment called
 CC notch extracellular truncation (NEXT). This fragment is then
 CC cleaved by presenilin dependent gamma-secretase to release a
 CC notch-derived peptide containing the intracellular domain (NICD)
 CC from the membrane (By similarity).
 CC -!- PTM: Phosphorylated (By similarity).
 CC -!- SIMILARITY: Belongs to the NOTCH family.
 CC -!- SIMILARITY: Contains 36 EGF-like domains.
 CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.
 CC -!- SIMILARITY: Contains 5 ANK repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X57405; CAA40667.1; -.
 CC HSPSP; P00740; 1EDM.
 CC InterPro; IPR002110; ANK.
 CC InterPro; IPR000152; Asx_hydroxyl_S.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR001881; EGF_Ca.
 CC InterPro; IPR001438; EGF_II.
 CC InterPro; IPR006209; EGF_Like.
 CC InterPro; IPR002049; Laminin_EGF.
 CC InterPro; IPR008297; Notch_dom.
 CC InterPro; IPR008000; Notch_dom.
 CC Pfam; PF000023; ank; 6.
 CC Pfam; PF00008; EGF; 35.
 CC Pfam; PF00066; notch; 3.

DR PIRSF; PIRSF002279; Notch; 1.
 DR PRINTS; PRO0010; EGFLOOD.
 DR PRINTS; PRO0011; EGFLAMININ.
 DR PRINTS; PRO1452; NOTCH.
 DR SMART; SM00248; ANK; 6.
 DR SMART; SM00179; EGF_CA; 25.
 DR SMART; SM00004; NU; 2.
 DR PROSITE; PS50297; ANK REP REGION; 1.
 DR PROSITE; PS50088; ANK REPEAT; 4.
 DR PROSITE; PS00010; ASX_HYDROXYL; 22.
 DR PROSITE; PS00022; EGF_1; 35.
 DR PROSITE; PS01186; EGF_2; 26.
 DR PROSITE; PS50026; EGF_3; 36.
 DR PROSITE; PS01187; EGF_CA; 21.
 DR Receptor; Transcription regulation; Activator; Differentiation;
 KW Developmental protein; Repeat; ANK repeat; EGF-like domain;
 KW Transmembrane; Glycoprotein; Signal; Phosphorylation.
 FT SIGNAL 1 18
 FT CHAIN 19 2531
 FT CHAIN 1711 2531
 FT CHAIN 1744 2531
 FT NOTCH INTRACELLULAR DOMAIN (BY
 FT SIMILARITY).
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT EGF-LIKE 1.
 FT EGF-LIKE 2.
 FT EGF-LIKE 3.
 FT EGF-LIKE 4.
 FT EGF-LIKE 5.
 FT EGF-LIKE 6.
 FT EGF-LIKE 7.
 FT EGF-LIKE 8.
 FT EGF-LIKE 9.
 FT EGF-LIKE 10.
 FT EGF-LIKE 11.
 FT EGF-LIKE 12.
 FT EGF-LIKE 13.
 FT EGF-LIKE 14.
 FT EGF-LIKE 15.
 FT EGF-LIKE 16.
 FT EGF-LIKE 17.
 FT EGF-LIKE 18.
 FT EGF-LIKE 19.
 FT EGF-LIKE 20.
 FT EGF-LIKE 21.
 FT EGF-LIKE 22.
 FT EGF-LIKE 23.
 FT EGF-LIKE 24.
 FT EGF-LIKE 25.
 FT EGF-LIKE 26.
 FT EGF-LIKE 27.
 FT EGF-LIKE 28.
 FT EGF-LIKE 29.
 FT EGF-LIKE 30.
 FT EGF-LIKE 31.
 FT EGF-LIKE 32.
 FT EGF-LIKE 33.
 FT EGF-LIKE 34.
 FT EGF-LIKE 35.
 FT EGF-LIKE 36.
 FT LIN/NOTCH 1.
 FT LIN/NOTCH 2.
 FT LIN/NOTCH 3.
 FT ANK 1.
 FT ANK 2.
 FT ANK 3.
 FT ANK 4.
 FT ANK 5.
 FT POLY-ALA.
 FT POLY-GLU.
 FT POLY-PRO.
 FT 2258

FT	DOMAIN	2497	2500	POLY-SER.	CLEAVAGE BY (FURIN-LIKE PROTEASE) (BY
FT	SITE	1654	1655	SIMILARITY)	
FT	DISULFID	24	37	BY SIMILARITY.	
FT	DISULFID	31	46	BY SIMILARITY.	
FT	DISULFID	43	57	BY SIMILARITY.	
FT	DISULFID	63	74	BY SIMILARITY.	
FT	DISULFID	68	87	BY SIMILARITY.	
FT	DISULFID	89	98	BY SIMILARITY.	
FT	DISULFID	106	117	BY SIMILARITY.	
FT	DISULFID	111	127	BY SIMILARITY.	
FT	DISULFID	129	138	BY SIMILARITY.	
FT	DISULFID	144	155	BY SIMILARITY.	
FT	DISULFID	149	164	BY SIMILARITY.	
FT	DISULFID	166	175	BY SIMILARITY.	
FT	DISULFID	182	195	BY SIMILARITY.	
FT	DISULFID	189	204	BY SIMILARITY.	
FT	DISULFID	206	215	BY SIMILARITY.	
FT	DISULFID	222	233	BY SIMILARITY.	
FT	DISULFID	227	243	BY SIMILARITY.	
Query Match 12.7%; Score 95.5; DB 1; Length 2531;					
Best Local Similarity 26.8%; Pred. No. 0.2;					
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7					
QY	15 ISACAMLLCHGSLQHTFQHHLRPEGTCE-VIAAHR	-----NGTCE	DTAGFTCR	CEGYHDP	TCLSEVNECNSN 723
DB	679 IDECAGSPCH				-----CNKN 56
QY	57 -----RIEERSQTVKSCLPKVKAGTTRNRPSVDASIVIGKWCCEMP	PCLEGE	CKTLP	111	
DB	724 PCIHGACRDLGNGYKDCAPG-WSGTNCD	-----INN	NECESN	PCVNGTCKOM-	771
QY	112 DNSGMMC 118				
DB	772 -TSGYVC 777				

RESULT 5

CRB_DROME

ID_CRB_DROME STANDARD; PRT: 2139 AA.

AC P10040;

AD 01-MAR-1989 (Rel. 10, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Crumbs protein precursor (95F).

GN CRB.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota, Metazoa, Arthropoda, Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RA [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Oregon-R; TISSUE=Embryo;

RX MEDLINE=90263104; PubMed=2344615;

RA Tepass U., Theres C., Knust E.;

RT "Crumbs encodes an EGF-like protein expressed on apical membranes of Drosophila epithelial cells and required for organization of epithelia.";

RL Cell 61:787-799(1990).

RN [2]

RP SEQUENCE OF 1663-1955 FROM N.A.

RC TISSUE=Embryo;

RA MEDLINE=87218537; PubMed=3107986;

RX Knust E., Dietrich U., Tepass U., Brenner K.A., Weigelt D., Vaeglin H., Campos-Ortega J.A.;

RT "EGF homologous sequences encoded in the genome of Drosophila melanogaster, and their relation to neurogenic genes.";

RL EMBO J. 6:761-766(1987).

CC -!- FUNCTION: May play a role in the development of epithelia, possibly for the establishment and/or maintenance of cell polarity. It may act as a signal.

CC

InterPro; IPR000152; Asx hydroxyl_s.
 DR InterPro; IPR001774; DSL.
 DR InterPro; IPR000742; EGF 2.
 DR InterPro; IPR001891; EGF_Ca.
 DR InterPro; IPR001438; EGF_II.
 DR InterPro; IPR006209; EGF_Like.
 DR InterPro; IPR002049; Laminin EGF.
 DR InterPro; IPR009041; PMP inhibitor.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF01414; DSL; 1.
 DR Pfam; PF00008; EGF; 14.
 DR PRINTS; PRO0010; EGFBLD.
 DR PRINTS; PRO0011; EGF_LAMININ.
 DR SMART; SM00051; DSL; 1.
 DR SMART; SM00179; EGF_CA; 10.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS00010; ASX HYDROXYL; 10.
 DR PROSITE; PS00022; EGF_1; 16.
 DR PROSITE; PS01186; EGF_2; 12.
 DR PROSITE; PS00026; EGF_3; 15.
 DR PROSITE; PS01187; EGF_CA; 8.
 DR PROSITE; PS01208; VWFC 1; FALSE NEG.
 KW Calcium-binding; EGF-like domain; Developmental protein;
 KW Repeat; Transmembrane; Signal.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 1213 JAGGED 3.
 FT DOMAIN 27 1064 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1065 1087 POTENTIAL.
 FT DOMAIN 1088 1213 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 164 226 DSL.
 FT DOMAIN 227 260 EGF-LIKE 1.
 FT DOMAIN 258 291 EGF-LIKE 2.
 FT DOMAIN 293 331 EGF-LIKE 3.
 FT DOMAIN 333 369 EGF-LIKE 4.
 FT DOMAIN 371 407 EGF-LIKE 5.
 FT DOMAIN 409 445 EGF-LIKE 6.
 FT DOMAIN 447 482 EGF-LIKE 7.
 FT DOMAIN 484 520 EGF-LIKE 8.
 FT DOMAIN 522 558 EGF-LIKE 9.
 FT DOMAIN 592 624 EGF-LIKE 10.
 FT DOMAIN 626 662 EGF-LIKE 11.
 FT DOMAIN 664 700 EGF-LIKE 12.
 FT DOMAIN 702 738 EGF-LIKE 13.
 FT DOMAIN 746 777 EGF-LIKE 14.
 FT DOMAIN 779 815 EGF-LIKE 15.
 FT DOMAIN 817 853 EGF-LIKE 16.
 FT DOMAIN 860 914 VWFC.
 FT DOMAIN 918 956 EGF-LIKE 17.
 FT DISULFID 231 242 BY SIMILARITY.
 FT DISULFID 235 248 BY SIMILARITY.
 FT DISULFID 250 259 BY SIMILARITY.
 FT DISULFID 262 273 BY SIMILARITY.
 FT DISULFID 268 279 BY SIMILARITY.
 FT DISULFID 281 290 BY SIMILARITY.
 FT DISULFID 297 309 BY SIMILARITY.
 FT DISULFID 303 319 BY SIMILARITY.
 FT DISULFID 321 330 BY SIMILARITY.
 FT DISULFID 337 348 BY SIMILARITY.
 FT DISULFID 342 357 BY SIMILARITY.
 FT DISULFID 359 368 BY SIMILARITY.
 FT DISULFID 375 386 BY SIMILARITY.
 FT DISULFID 380 395 BY SIMILARITY.
 FT DISULFID 397 406 BY SIMILARITY.
 FT DISULFID 413 424 BY SIMILARITY.
 FT DISULFID 418 433 BY SIMILARITY.
 FT DISULFID 435 444 BY SIMILARITY.
 FT DISULFID 451 461 BY SIMILARITY.
 FT DISULFID 455 470 BY SIMILARITY.
 FT DISULFID 472 481 BY SIMILARITY.
 FT DISULFID 488 499 BY SIMILARITY.
 FT DISULFID 493 508 BY SIMILARITY.
 FT DISULFID 510 519 BY SIMILARITY.
 FT DISULFID 526 537 BY SIMILARITY.

FT DISULFID 531 546 BY SIMILARITY.
 FT DISULFID 548 557 BY SIMILARITY.
 FT DISULFID 630 641 BY SIMILARITY.
 FT DISULFID 635 650 BY SIMILARITY.
 FT DISULFID 652 661 BY SIMILARITY.
 FT DISULFID 668 679 BY SIMILARITY.
 FT DISULFID 673 688 BY SIMILARITY.
 FT DISULFID 690 699 BY SIMILARITY.
 FT DISULFID 706 717 BY SIMILARITY.
 FT DISULFID 711 726 BY SIMILARITY.
 FT DISULFID 728 737 BY SIMILARITY.
 FT DISULFID 783 794 BY SIMILARITY.
 FT DISULFID 788 803 BY SIMILARITY.
 FT DISULFID 805 814 BY SIMILARITY.
 FT DISULFID 821 832 BY SIMILARITY.
 FT DISULFID 826 841 BY SIMILARITY.
 FT DISULFID 843 852 BY SIMILARITY.
 FT DOMAIN 938 941 POLY-PRO.
 FT CARBOHYD 119 139 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 556 556 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 742 742 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 957 957 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 988 988 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1042 1042 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1213 AA; 133365 MW; 5C5F16A7E20D9534 CRC64;
 Query Match 12.1%; Score 91.5; DB 1; Length 1213;
 Best Local Similarity 24.8%; Pred. No. 0.23;
 Matches 35; Conservative 19; Mismatches 50; Indels 37; Gaps 8;
 QY 17 ACAMLLC--HGSLOHTFQOHLHRP---EGGTCEV---IAAHRCCNKNRIBERSQTVKC 67
 DB 336 ACLSNPCANGTKETSQSYECHECAIGWSGTSCINVDCTPNQCKHGTGQDLVNGFKC 395
 QY 68 SCLP---GKV---AGTTRNRPSCVDASI---VIGKWCMEP-----100
 DB 396 ACPPHWGTGTCQIDANECEDEXP-CVNAKSCHNLIGAYFCECLPFGWSGQNCINNDCKGQ 454
 QY 101 CLEGECKTLPDNGWMCATG 121
 DB 455 CLNGGTCKDLVNGYRCLCPPG 475
 RESULT 7
 DP87_DICDI STANDARD; PRT; 555 AA.
 AC Q04503;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Prespore protein DP87 precursor.
 GN COTD OR DP87.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX3;
 RX MEDLINE=94008533; PubMed=8404532;
 RA Ozaki T., Nakao H., Orii H., Morio T., Takeuchi I., Tasaka M.;
 RT "Developmental regulation of transcription of a novel
 prespore-specific gene (Op87) in Dictyostelium discoideum.";
 RL Development 117:1299-1308(1993).
 CC -|- SUBCELLULAR LOCATION: Stored in prespore vacuoles until it is
 CC discharged into the interspace of spores during spore formation.
 CC -|- INDUCTION: By exogenous cAMP, repressed by DIP.
 CC -|- SIMILARITY: Contains 7 prespore motif repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D13973; BAA03083.1; --
 DR DictyBase; DDB0001389; --
 DR InterPro; IPR007643; Dict_spore_N.
 DR InterPro; IPR003645; FOLN.
 DR Pfam; PF04562; Dict_spore_N; 1.
 DR SMART; SM00274; FOLN; 6.
 KW Glycoprotein; Sporulation; Signal; Repeat.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 555 PRESPORE PROTEIN DP87.
 FT REPEAT 149 161 PRESPORE MOTIF 1.
 FT REPEAT 176 188 PRESPORE MOTIF 2.
 FT REPEAT 210 222 PRESPORE MOTIF 3.
 FT REPEAT 241 253 PRESPORE MOTIF 4.
 FT REPEAT 305 317 PRESPORE MOTIF 5.
 FT REPEAT 333 345 PRESPORE MOTIF 6.
 FT REPEAT 363 375 PRESPORE MOTIF 7.
 FT DOMAIN 438 473 THR-RICH.
 FT DOMAIN 476 555 ALA/SER-RICH.
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 555 AA; 58673 MW; 6DA8A6D1C36F8DD CRC64;

Query Match 12.1%; Score 91; DB 1; Length 555;
 Best Local Similarity 22.9%; Pred. No. 0.12;
 Matches 27; Conservative 14; Mismatches 53; Indels 24; Gaps 5;

QY 34 HHLHR-----PEGGTCEVIAHRCNKNRIERSQTV-----KCSCLPGKVAGTRNRPS 84
 DB 238 HNLGRFGCGEGHCEVLEKHPVCRNHPHPPPPPQICGNSVCGGYCTTINHGHTC 297
 QY 85 VDAISIVIGKWCMEPCLEGECKTLP-----DNSGW-MCATGNKIKTRIHP 131
 DB 298 IRGD----GYLCNQTRCPHDYQDETISTNVKSPKNDCKWRCPPGSSCFNSRNGP 351

RESULT 8

NOTC DROME STANDARD; PRT; 2703 AA.
 AC P07207; O97458; P04154; O9W478;
 DT 01-NOV-1986 (Rel. 03, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Neurogenic locus Notch protein precursor.
 GN N OR EG:140G11.1 OR EG:163A10.2 OR CG3936.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R; TISSUE=Embryo;
 RX MEDLINE=86079539; PubMed=3935325;
 RA Wharton K.A., Johansen K.M., Xu T., Artavanis-Tsakonas S.;
 RT "Nucleotide sequence from the neurogenic locus notch implies a gene
 RT product that shares homology with proteins containing EGF-like
 RT repeats.";
 RL Cell 43:567-581(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S, and Oregon-R; TISSUE=Embryo;
 RX MEDLINE=87064624; PubMed=3097517;
 RA Kidd S., Kelley M.R., Young M.W.;
 RT "Sequence of the notch locus of Drosophila melanogaster: relationship
 RT of the encoded protein to mammalian clotting and growth factors.";
 RL Mol. Cell. Biol. 6:3094-3108(1986).
 RN [3]
 RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champ M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Finnkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale B., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R;
 RX MEDLINE=20196011; PubMed=10731137;
 RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
 RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E.,
 RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,
 RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
 RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
 RA Modolell J., Peter A., Schoettler P., Werner M., Mourikioti F.,
 RA Beinart N., Dowe G., Schaefer U., Jaekle H., Bucheton A.,
 RA Callister D.M., Campbell L.A., Darlamitson A., Henderson N.S.,
 RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
 RA Glover D.M.;
 RT "From sequence to chromosome: the tip of the X chromosome of D.
 RT melanogaster.";
 RL Science 287:2220-2222 (2000).
 RN [5]
 RP SEQUENCE OF 2505-2611 FROM N.A.
 RX MEDLINE=85099329; PubMed=2981631;
 RA Wharton K.A., Vedvobnick B., Finnerty V.G., Artavanis-Tsakonas S.;
 RT "opa: a novel family of transcribed repeats shared by the Notch locus
 RT and other developmentally regulated loci in D. melanogaster.";
 RL Cell 40:55-62(1985).
 RN [6]
 RP SEQUENCE OF 1-8 FROM N.A.
 RX MEDLINE=87257846; PubMed=3037327;
 RA Kelley M.R., Kidd S., Berg R.L., Young M.W.;
 RT "Restriction of P-element insertions at the Notch locus of Drosophila
 RT melanogaster.";
 RL Mol. Cell. Biol. 7:1545-1548(1987).
 RN [7]
 RP INTERACTION WITH DX, AND MUTANT SU42C.

RX MEDLINE=94215489; PubMed=8162848;
RA Diederich R.J., Matsumo K., Hing H., Artavanis-Tsakonas S.;
RT "Cytoolic interaction between deltex and Notch ankyrin repeats
RT implicates deltex in the Notch signaling pathway.";
RL Development 120:473-481(1994).
RN [8]
RN INTERACTION WITH DX.
RX MEDLINE=95401878; PubMed=7671825;
RA Matsumo K., Diederich R.J., Go M.J., Blaumueller C.M.,
RA Artavanis-Tsakonas S.;
RT "Deltex acts as a positive regulator of Notch signaling through
RT interactions with the Notch ankyrin repeats.";
RL Development 121:2633-2644(1995).
RN [9]
RN S3 CLEAVAGE BY PSN.
RX MEDLINE=99221487; PubMed=10206646;
RA Struhl G., Greenwald I.;
RT "Presenilin is required for activity and nuclear access of Notch in
RT Drosophila.";
RL Nature 398:522-525(1999).
RN [10]
RN S3 CLEAVAGE BY PSN.
RX MEDLINE=99221488; PubMed=10206647;
RA Ye Y., Lukinova N., Fortini M.E.;
RT "Neurogenic phenotypes and altered Notch processing in Drosophila
RT Presenilin mutants.";
RL Nature 398:525-529(1999).
RN [11]
RN S2 CLEAVAGE BY KUZ.
RX MEDLINE=21657146; PubMed=11799064;
RA Lieber T., Kidd S., Young M.W.;
RT "kuzbanian-mediated cleavage of Drosophila Notch.";
RL Genes Dev. 16:209-221(2002).
RN [12]
RN MUTANT MCDs.
RX MEDLINE=21575956; PubMed=11719214;
RA Raman P., Khechumian K., Seugnet L., Arbogast N., Ackermann C.,
RA Heitzler P.;
RT "Novel Notch alleles reveal a Deltex-dependent pathway repressing
RT neural fate.";
RL Curr. Biol. 11:1729-1738(2001).
RN [13]
RN REVIEW.
RX MEDLINE=22256570; PubMed=12369105;
RA Fortin P.;
RT "General outlines of the molecular genetics of the Notch signalling
RT pathway in Drosophila melanogaster: a review.";
RL Hereditas 136:89-96(2002).
CC -!- FUNCTION: Signaling protein, which regulates, with both positive
CC and negative signals, the differentiation of at least central and
CC peripheral nervous system and eye, wing disk, oogenesis, segmental
CC appendages such as antennae and legs, and muscles, through lateral
CC inhibition or induction. Functions as a receptor for membrane-
CC bound ligands Delta and Serrate to regulate cell-fate
CC determination. Upon ligand activation, and releasing from the cell
CC membrane, the Notch intracellular domain (NICD) forms a
CC transcriptional activator complex with Su(H) (Suppressor of
CC hairless) and activates genes of the E(spl) complex. Essential for
CC proper differentiation of ectoderm.
CC -!- SUBUNIT: Interacts with Su(H) when activated. Interacts with Dx
CC via its ANK repeats.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Upon activation and
CC S3 cleavage, it is released from the cell membrane and enters into
CC the nucleus in conjunction with Su(H).
CC -!- PFM: Upon binding its ligands such as Delta or Serrate, it is
CC cleaved (S2 cleavage) in its extracellular domain, close to the
CC transmembrane domain. S2 cleavage is probably mediated by Kuz. It
CC is then cleaved (S3 cleavage) downstream of its transmembrane
CC domain, releasing it from the cell membrane. S3 cleavage requires
CC Psn.
CC -!- SIMILARITY: Belongs to the NOTCH family.
CC -!- SIMILARITY: Contains 36 EGF-like domains.
CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.

CC -!- SIMILARITY: Contains 6 ANK repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16152; AAB59220.1; -;
DR EMBL; M16153; AAB59220.1; JOINED.
DR EMBL; M16149; AAB59220.1; JOINED.
DR EMBL; M16150; AAB59220.1; JOINED.
DR EMBL; M16151; AAB59220.1; JOINED.
DR EMBL; K03508; AAA28725.1; -;
DR EMBL; M13689; AAA28725.1; JOINED.
DR EMBL; K03507; AAA28725.1; JOINED.
DR EMBL; AR003426; AAF45848.2; -;
DR EMBL; AL035436; CAB37610.1; -;
DR EMBL; AL035395; CAB37610.1; JOINED.
DR EMBL; M12175; AAA74496.1; -;
DR EMBL; M16025; AAA28726.1; -;
CC -----
CC Query Match 11.5%; Score 86.5; DB 1; Length 2703;
CC Best Local Similarity 25.9%; Pred. No. 1.7;
CC Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;
CC -----
CC QY 40 EGGTCBEV----IAAHRCCNKNRIERSQTVKCSCLPGKVA-----GTTNRPS 83
CC 1214 QQONCELNIDDCAPNCPQNGGTCRDVNMVNFSCPPGTGIIICEINKDKCKPGACINNGS 1273
CC QY 84 CVDASIVIGKWCMEPEPCLEGBECK 108
CC 1274 CIDR---VGGFEVCVQPGFVGARCE 1295
CC -----
CC RESULT 9
CC DLL4_HUMAN
CC ID DLL4_HUMAN STANDARD; PRT; 685 AA.
CC AC Q9NR61; Q9NOT9;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Delta-like protein 4 precursor (Drosophila Delta homolog 4)
CC DE (UNQ1895/PRO4341).
CC GN DLL4.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Brain;
CC RX MEDLINE=20296719; PubMed=10837024;
CC RA Shutter J.R., Scully S., Pan W., Richards W.G., Kitajewski J.,
CC Deblandre G.A., Kintner C.R., Stark K.B.;
CC RT "Dl14, a novel Notch ligand expressed in arterial endothelium.";
CC RL Genes Dev. 14:1313-1318(2000).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RA Sakano S.;
CC RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC RN [3]
CC RP SEQUENCE FROM N.A.
CC RA Yoneya T., Tahara T., Nagao K., Yamada Y., Yamamoto T., Miyatani S.,
CC Nishikawa M.;
CC RT "Molecular cloning of Delta-4, a new mouse and human Notch ligand.";
CC RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC RN [4]
CC RP SEQUENCE FROM N.A.
CC RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,

RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heidens S.,
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
 RA Vandlen R., Watanabe C., Wiand D., Woods K., Xie M.-H., Yansura D.,
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
 RA Godowski P.,
 RT "The secreted protein discovery initiative (SPDI), a large-scale
 RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment.";
 RL Genome Res. 13:2265-2270(2003).
 RN [5]
 RP SEQUENCE OF 33-685 FROM N.A.
 RC
 RA TISSUE=Placenta;
 RA Malmos C., Modlich U., Lewis J., Harris A., Bicknell R.,
 RA Ish-Horowicz D.;
 RT "A novel Delta gene expressed in embryonic and tumour vasculature.";
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RL
 CC - FUNCTION: Plays a role in the Notch signaling pathway. Activates
 CC - Notch-1 and Notch-4 (By similarity).
 CC - SUBUNIT: Binds to Notch-1 and Notch-4 (By similarity).
 CC - SUBCELLULAR LOCATION: Type I membrane protein (Probable).
 CC - TISSUE SPECIFICITY: Expressed in vascular endothelium.
 CC - DOMAIN: The Delta-Serrate-Lag2 (DSL) domain is required for
 CC binding to the Notch receptor.
 CC - SIMILARITY: Contains 8 EGF-like domains.
 CC - SIMILARITY: Contains 1 DSL domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF253468; AAF76427.1; -
 CC EMBL: AB036931; BAB16085.1; -
 CC EMBL: AB043894; BAB18581.1; -
 CC EMBL: AY358894; AAK89253.1; -
 CC EMBL: AF279305; AAF81912.1; -
 CC FIR; JC7570; JC7570.
 CC HSP; P00740; LEDM.
 CC Genew; HGNC:2910; DLL4.
 CC MIM: 605185; -
 CC GO; GO:0005112; F:Notch binding; TAS.
 CC GO; GO:0006015; P:circulation; TAS.
 CC GO; GO:0007165; P:signal transduction; TAS.
 CC InterPro; IPR000152; Asx_hydroxyl_s.
 CC InterPro; IPR001774; DSL.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR001881; EGF_Ca.
 CC InterPro; IPR001438; EGF_I1.
 CC InterPro; IPR00209; EGF_like.
 CC InterPro; IPR002049; Laminin_EGF.
 CC Pfam; PF01414; DSL; 1.
 CC Pfam; PF00008; EGF; 7.
 CC PRINTS; PR00010; EGFBL00D.
 CC PRINTS; PR00011; EGFFLAMIN.
 CC SMART; SM00051; DSL; 1.
 CC SMART; SM00179; EGF_CA; 2.
 CC PROSITE; PS00010; ASX_HYDROXYL; 1.
 CC PROSITE; PS00022; EGF_1; 8.
 CC PROSITE; PS01186; EGF_2; 7.
 CC PROSITE; PS00026; EGF_3; 8.
 CC Signal; EGF-like domain; Repeat; Transmembrane; Developmental protein;
 CC Differentiation; Glycoprotein.
 KW
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 685 DELTA-LIKE PROTEIN 4.
 FT DOMAIN 27 529 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 530 550 POTENTIAL.
 FT DOMAIN 551 685 CYTOPLASMIC (POTENTIAL).
 FT

FT DOMAIN 155 217 DSL.
 FT DOMAIN 218 251 EGF-LIKE 1.
 FT DOMAIN 252 282 EGF-LIKE 2.
 FT DOMAIN 284 322 EGF-LIKE 3.
 FT DOMAIN 324 360 EGF-LIKE 4.
 FT DOMAIN 362 400 EGF-LIKE 5.
 FT DOMAIN 402 438 EGF-LIKE 6.
 FT DOMAIN 440 476 EGF-LIKE 7.
 FT DOMAIN 480 518 EGF-LIKE 8.
 FT DISULFID 222 233 BY SIMILARITY.
 FT DISULFID 226 239 BY SIMILARITY.
 FT DISULFID 241 250 BY SIMILARITY.
 FT DISULFID 253 254 BY SIMILARITY.
 FT DISULFID 259 270 BY SIMILARITY.
 FT DISULFID 272 281 BY SIMILARITY.
 FT DISULFID 288 300 BY SIMILARITY.
 FT DISULFID 294 310 BY SIMILARITY.
 FT DISULFID 312 321 BY SIMILARITY.
 FT DISULFID 328 339 BY SIMILARITY.
 FT DISULFID 333 348 BY SIMILARITY.
 FT DISULFID 350 359 BY SIMILARITY.
 FT DISULFID 366 377 BY SIMILARITY.
 FT DISULFID 371 388 BY SIMILARITY.
 FT DISULFID 390 399 BY SIMILARITY.
 FT DISULFID 406 417 BY SIMILARITY.
 FT DISULFID 411 426 BY SIMILARITY.
 FT DISULFID 428 437 BY SIMILARITY.
 FT DISULFID 444 464 BY SIMILARITY.
 FT DISULFID 466 475 BY SIMILARITY.
 FT DISULFID 484 495 BY SIMILARITY.
 FT DISULFID 489 506 BY SIMILARITY.
 FT DISULFID 508 517 BY SIMILARITY.
 FT CARBOHYD 188 198 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 183 193 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 393 393 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 685 AA; 74604 MW; 6CF89D3C220ACC89 CRC64;
 Query Match 11.4%; Score 86; DB 1; Length 685;
 Best Local Similarity 25.0%; Pred. No. 0.45;
 Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;
 QY 14 WISAC-----AMLLCHGSLQHTFQOHLHPEGTCEVIAHRCCKNKRTEERSQVVKSC 69
 DB 268 WQCTCDEGWGGLFCDDQLNYC--THSPCKNGATC-----SNSGQSYT--CTC 312
 QY 70 LFGKVAGTTRNRPSCVDASIVIGRWKMEPCLEGECKTLPDNGWMCATG 121
 DB 313 RPYGTG-----VDCELELSE--CDSNPNCRNGSGCKQEDGYHCLCPPG 353
 RESULT 10
 ID JAG1_HUMAN STANDARD; PRT; 1218 AA.
 AC P78504; O14902; Q15122; Q15816;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Jagged1 precursor (Jagged1) (hJ1).
 OS JAG1 OR JAG1.
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97422615; PubMed=9268641;
 RA Oda T., Elkahloune A.G., Meltzer P.S., Chandrasekharappa S.C.;
 RT "Identification and cloning of the human homolog (JAG1) of the rat
 RT Jagged1 gene from the Alagille syndrome critical region at 20p12.";
 RL Genomics 43:376-379(1997).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT AGS CVS-184.

- TISSUE=Bone marrow;
 RX MEDLINE=97351505; PubMed=9207788;
 RA Oda T., Elkhoun A.G., Pike B.L., Okajima K., Krantz I.D., Genin A.,
 RA Piccoli D.A., Meltzer P.S., Spinner N.B., Collins F.S.,
 RA Chandrasekharappa S.C.;
 RT "Mutations in the human Jagged1 gene are responsible for Alagille
 syndrome.";
 RL Nat. Genet. 16:235-242(1997).
 RN [9]
 RP DEVELOPMENTAL STAGE.
 RX MEDLINE=20436345; PubMed=10978356;
 RA Jones E.A., Clement-Jones M., Wilson D.I.;
 RT "JAGGED1 expression in human embryos: correlation with the Alagille
 syndrome phenotype.";
 RL J. Med. Genet. 37:663-668(2000).
 RN [10]
 RP VARIANTS AGS CYS-184 AND HIS-184.
 RX MEDLINE=98254456; PubMed=9585603;
 RA Krantz I.D., Colliton R.P., Genin A., Rand E.B., Li L., Piccoli D.A.,
 RA Spinner N.B.;
 RT "Spectrum and frequency of jagged1 (JAG1) mutations in Alagille
 syndrome patients and their families.";
 RL Am. J. Hum. Genet. 62:1361-1369(1998).
 RN [11]
 RP VARIANTS AGS HIS-79; THR-127; ARG-129; LEU-163; GLY-184; SER-187;
 RX MEDLINE=99238888; PubMed=10220506;
 RA Crosnier C., Driancourt C., Raynaud N., Dhorne-Pollet S., Pollet N.,
 RA Bernard O., Hachouel M., Meunier-Rotival M.;
 RT "Mutations in JAGGED1 gene are predominantly sporadic in Alagille
 syndrome.";
 RL Gastroenterology 116:1141-1148(1999).
 RN [12]
 RP VARIANTS AGS THR-152 AND LEU-184.
 RX MEDLINE=20004539; PubMed=10533065;
 RA Pilia G., Uda M., Macis D., Frau F., Crisponi L., Balli F.,
 RA Barbera C., Colombo C., Frediani T., Gatti R., Iorio R., Marazzi M.G.,
 RA Marcellini M., Musumeci S., Nebbia G., Vajro P., Ruffa G., Zancan L.,
 RA Cao A., DeVirgili S.;
 RT "Jagged-1 mutation analysis in Italian Alagille syndrome patients.";
 RL Hum. Mutat. 14:394-400(1999).
 RN [13]
 RP VARIANTS AGS TYR-229 AND ARG-386.
 RX MEDLINE=20514559; PubMed=11058998;
 RA Heritage M.L., MacMillan J.C., Colliton R.P., Genin A., Spinner N.B.,
 RA Anderson G.J.;
 RT "Jagged1 (JAG1) mutation detection in an Australian Alagille syndrome
 population.";
 RL Hum. Mutat. 16:408-416(2000).
 RN [14]
 RP VARIANT TOF ASP-274.
 RX MEDLINE=21067871; PubMed=11152664;
 RA Eladad Z.A., Hamosh A., Biery N.J., Montgomery R.A., Duke M.,
 RA Elkins R., Dietz H.C.;
 RT "Familial Tetralogy of Fallot caused by mutation in the jagged1
 gene.";
 RL Hum. Mol. Genet. 10:163-169(2001).
 RN [15]
 RP VARIANTS AGS SER-37.
 RX MEDLINE=21096916; PubMed=11157803;
 RA Morrisette J.D., Colliton R.P., Spinner N.B.;
 RT "Defective intracellular transport and processing of JAG1 missense
 mutations in Alagille syndrome.";
 RL Hum. Mol. Genet. 10:405-413(2001).
 RN [16]
 RP VARIANTS AGS PHE-220 AND ARG-753.
 RX MEDLINE=20579880; PubMed=11139247;
 RA Crosnier C., Driancourt C., Raynaud N., Hachouel M.,
 RA Meunier-Rotival M.;
 RT "Fifteen novel mutations in the JAGGED1 gene of patients with Alagille
 syndrome.";
 RL Hum. Mutat. 17:72-73(2001).
 RN CC -!- FUNCTION: Ligand for multiple Notch receptors and involved in the
 CC mediation of Notch signaling. May be involved in cell-fate
 CC
- TISSUE=Bone marrow;
 RX MEDLINE=97351505; PubMed=9207788;
 RA Li L., Krantz I.D., Deng Y., Genin A., Banta A.B., Collins C.C.,
 RA Qi M., Trask B.J., Kuo W.L., Cochran J., Costa T., Pierpont M.B.M.,
 RA Rand E.B., Piccoli D.A., Hood L., Spinner N.B.;
 RT "Alagille syndrome is caused by mutations in human Jagged1, which
 encodes a ligand for Notch1.";
 RL Nat. Genet. 16:243-251(1997).
 RN [3]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RX MEDLINE=98122342; PubMed=9462510;
 RA Li L., Milner L.A., Deng Y., Iwata M., Banta A.B., Graf L.,
 RA Marcovina S., Friedman C., Trask B.J., Hood L., Torok-Storb B.;
 RT "The human homolog of rat Jagged1 expressed by marrow stroma inhibits
 differentiation of 32D cells through interaction with Notch1.";
 RL Immunity 8:43-55(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99180765; PubMed=10079256;
 RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,
 RA Banks A., Leiman J., Ward D., Ieh-Horowitz D., Artavanis-Teakonas S.;
 RT "Human ligands of the Notch receptor.";
 RL Am. J. Pathol. 154:785-794(1999).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99180765; PubMed=10079256;
 RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,
 RA Banks A., Leiman J., Ward D., Ieh-Horowitz D., Artavanis-Teakonas S.;
 RT "Human ligands of the Notch receptor.";
 RL Am. J. Pathol. 154:785-794(1999).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Scavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Leiva-Saiaho M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.B., McConachie L.J., McLay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Swann R.M., Sycamore N., Soderlund C., Steward C.A., Sulston J.E.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Thorpe A.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [7]
 RP SEQUENCE OF 14-1227 FROM N.A.
 RX MEDLINE=97115768; PubMed=8955070;
 RA Zimin A.B., Pepper M.S., McMahon G.A., Nguyen F., Montesano R.,
 RA Maciag T.;
 RT "An antisense oligonucleotide to the notch ligand jagged enhances
 fibroblast growth factor-induced angiogenesis in vitro.";
 RL J. Biol. Chem. 271:32499-32502(1996).
 RN [8]
 RP DISEASE.

CC decisions during hematopoiesis. Seems to be involved in early
 CC and late stages of mammalian cardiovascular development. Inhibits
 CC myoblast differentiation (By similarity). Enhances fibroblast
 CC growth factor-induced angiogenesis (in vitro).
 CC -!- SUBUNIT: Interacts with NOTCH1, NOTCH2 AND NOTCH3 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Widely expressed in adult and fetal tissues.
 CC in cervix epithelium expressed in undifferentiated subcolumnar
 CC reserve cells and squamous metaplasia. Expression is up-regulated
 CC in cervical squamous cell carcinoma. Expressed in bone marrow cell
 CC line HS-27a which supports the long-term maintenance of immature
 CC progenitor cells.
 CC -!- DEVELOPMENTAL STAGE: Expressed in 32-52 days embryos in the distal
 CC cardiac outflow tract and pulmonary artery, major arteries, portal
 CC vein, optic vesicle, otocyst, branchial arches, metanephros, and in
 CC pancreas, mesocardium, around the major bronchial branches, and in
 CC the neural tube.
 CC -!- DISEASE: Defects in JAG1 are the cause of Alagille syndrome (AGS)
 CC [MIM:118450]. AGS is an autosomal dominant developmental disorder
 CC that affects structures in the liver, heart, skeleton, eye,
 CC kidney and other organs.
 CC -!- DISEASE: Defects in JAG1 are a cause of tetralogy of Fallot
 CC (TOF) [MIM:187500]. TOF is a congenital heart anomaly which
 CC consists of pulmonary stenosis, ventricular septal defect,
 CC dextroposition of the aorta (aorta is on the right side instead of
 CC the left) and hypertrophy of the right ventricle. This condition
 CC results in a blue baby at birth due to inadequate oxygenation.
 CC Surgical correction is emergent.
 CC -!- SIMILARITY: Contains 15 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 DSL domain.
 CC -!- CAUTION: Ref.7 sequence differs from that shown due to a
 CC frameshift in position 1187.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC
 CC Query Match 11.4%; Score 86; DB 1; Length 1218;
 CC Best Local Similarity 23.8%; Pred.No. 0.81;
 CC Matches 34; Conservative 11; Mismatches 34; Indels 64; Gaps 9;
 CC
 CC QY 37 HRP-EGTCE-----VIAHRC-----C-NKNRIERSQTVKSC 69
 CC Db 303 HQPLNGTSGNTGPDYKQSCPEGYGPNCEIAHACLDPCHNRGSKETSIGFEC 362
 CC
 CC QY 70 LPQKVGATGTRNR-----PSCVDASIVIG-----KWW-----CEMEP 100
 CC Db 363 SPGWGTGTCSTNIDDCSPNNCSHGTTQCD--LVNGFKVCVPPWGTGKTCQLDANECEAKP 420
 CC QY 101 CLSGEECKTLPDN-----SGWM 117
 CC Db 421 CVNAKCKNLIASYCYDCLPGWM 443
 CC
 CC RESULT 11
 CC JAG1_MOUSE
 CC ID JAG1_MOUSE STANDARD; PRT; 1218 AA.
 CC AC Q9QXXO;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
 CC DE Jagged 1 precursor (Jagged1).
 CC GN JAG1.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_TaxID:10090;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., AND RECEPTOR INTERACTION.
 CC RC STRAIN=Swiss Webster / NIH;
 CC RX MEDLINE=20020271; PubMed=10551863;
 CC RA Shimizu K., Chiba S., Kumano K., Hosoya N., Takahashi T., Kanda Y.,
 CC RA Hamada Y., Yazaki Y., Hirai H.;
 CC RT "Mouse Jagged1 physically interacts with Notch2 and other Notch
 CC RT receptors: assessment by quantitative methods.";

RL J. Biol. Chem. 274:32961-32969 (1999).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.J., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udman T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy M., Hellon E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting J., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RN TISSUE SPECIFICITY.
 RX MEDLINE=20025753; PubMed=10556292;
 RA Loomes K.M., Underkofer L.A., Morabito J., Gottlieb S.,
 RA Piccoli D.A., Spinner N.B., Baldwin H.S., Oakey R.J.;
 RT "The expression of Jagged1 in the developing mammalian heart
 CC correlates with cardiovascular disease in Alagille syndrome.";
 CC Hum. Mol. Genet. 8:2443-2449 (1999).
 CC -!- FUNCTION: Ligand for multiple Notch receptors and involved in the
 CC mediation of Notch signaling. May be involved in cell-fate
 CC decisions during hematopoiesis. Seems to be involved in early
 CC and late stages of mammalian cardiovascular development. Inhibits
 CC myoblast differentiation (By similarity). May regulate fibroblast
 CC growth factor-induced angiogenesis.
 CC -!- SUBUNIT: Interacts with NOTCH1, NOTCH2 AND NOTCH3.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Widely expressed in many tissues, with highest
 CC expression in brain, heart, muscle and thymus.
 CC -!- DEVELOPMENTAL STAGE: At 8.75-9.75 dpc expression was detected
 CC in structures that include those destined to contribute to the
 CC cardiovascular system of the adult heart. Expression was also
 CC detected in the mesencephalon and rhombencephalon.
 CC -!- DOMAIN: The DSL domain is indispensable and sufficient for binding
 CC to NOTCH2.
 CC -!- SIMILARITY: Contains 15 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 DSL domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF171092; AAF15505.1; -;
 CC EMBL; BC058675; AAH58675.1; -;
 CC HSPB; P00740; 1EDM.
 CC MGD; MGI:1095416; Jag1.
 CC GO; GO:0005576; C:extracellular; ISS.
 CC GO; GO:0005509; F:calcium ion binding; NAS.
 CC GO; GO:0008083; F:growth factor activity; ISS.
 CC GO; GO:0005112; F:Notch binding; IPI.
 CC GO; GO:0005198; F:structural molecule activity; ISS.
 CC GO; GO:0001525; P:angiogenesis; ISS.
 CC GO; GO:0001709; P:cell fate determination; ISS.
 CC GO; GO:0045446; P:endothelial cell differentiation; ISS.
 CC GO; GO:0030097; P:hematopoiesis; ISS.

DR GO: 0030216; P:keratinocyte differentiation; ISS.
 DR GO: 0002011; P:morphogenesis of an epithelial sheet; IMP.
 DR GO: 0045445; P:myoblast differentiation; ISS.
 DR GO: 0007219; P:N signaling pathway; ISS.
 DR GO: 0007399; P:neurogenesis; ISS.
 DR GO: 0042127; P:regulation of cell proliferation; ISS.
 DR InterPro: IPR000152; Asx_hydroxyl_S.
 DR InterPro: IPR001774; DSL.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_CA.
 DR InterPro: IPR001438; EGF_II.
 DR InterPro: IPR006209; EGF_Like.
 DR InterPro: IPR001007; WVF_C.
 DR Pfam: PF01444; DSL; 1.
 DR Pfam: PF00008; EGF; 14.
 DR PRINTS: PR00010; EGFBL00D.
 DR SMART: SMC00511; DSL; 1.
 DR SMART: SMC0179; EGF_CA; 10.
 DR SMART: SMC0214; VWC; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 10.
 DR PROSITE: PS00022; EGF_1; 16.
 DR PROSITE: PS01186; EGF_2; 12.
 DR PROSITE: PS00026; EGF_3; 15.
 DR PROSITE: PS01187; EGF_CA; 8.
 DR Calcium-binding; EGF-Like domain; Glycoprotein; Developmental protein;
 KW Repeat; Transmembrane; Signal.
 FT SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 1218 JAGGED 1.
 FT DOMAIN 34 1067 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1068 1093 POTENTIAL.
 FT DOMAIN 1094 1218 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 167 229 DSL.
 FT DOMAIN 230 266 EGF-LIKE 1 (ATYPICAL).
 FT DOMAIN 296 334 EGF-LIKE 2.
 FT DOMAIN 336 372 EGF-LIKE 3.
 FT DOMAIN 374 410 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 412 448 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 450 485 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 487 523 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 525 561 EGF-LIKE 8.
 FT DOMAIN 574 627 EGF-LIKE 9.
 FT DOMAIN 629 665 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 667 703 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 705 741 EGF-LIKE 12.
 FT DOMAIN 744 780 EGF-LIKE 13.
 FT DOMAIN 782 818 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 820 856 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
 FT DISULFID 300 312 BY SIMILARITY.
 FT DISULFID 306 322 BY SIMILARITY.
 FT DISULFID 324 333 BY SIMILARITY.
 FT DISULFID 340 351 BY SIMILARITY.
 FT DISULFID 345 360 BY SIMILARITY.
 FT DISULFID 362 371 BY SIMILARITY.
 FT DISULFID 378 389 BY SIMILARITY.
 FT DISULFID 383 398 BY SIMILARITY.
 FT DISULFID 400 409 BY SIMILARITY.
 FT DISULFID 416 427 BY SIMILARITY.
 FT DISULFID 421 436 BY SIMILARITY.
 FT DISULFID 438 447 BY SIMILARITY.
 FT DISULFID 454 464 BY SIMILARITY.
 FT DISULFID 458 473 BY SIMILARITY.
 FT DISULFID 475 484 BY SIMILARITY.
 FT DISULFID 491 502 BY SIMILARITY.
 FT DISULFID 496 511 BY SIMILARITY.
 FT DISULFID 513 522 BY SIMILARITY.
 FT DISULFID 529 540 BY SIMILARITY.
 FT DISULFID 534 549 BY SIMILARITY.
 FT DISULFID 551 560 BY SIMILARITY.
 FT DISULFID 578 605 BY SIMILARITY.
 FT DISULFID 599 615 BY SIMILARITY.
 FT DISULFID 617 626 BY SIMILARITY.
 FT DISULFID 633 644 BY SIMILARITY.
 FT DISULFID 638 653 BY SIMILARITY.

FT DISULFID 655 664 BY SIMILARITY.
 FT DISULFID 671 682 BY SIMILARITY.
 FT DISULFID 676 691 BY SIMILARITY.
 FT DISULFID 693 702 BY SIMILARITY.
 FT DISULFID 709 720 BY SIMILARITY.
 FT DISULFID 714 729 BY SIMILARITY.
 FT DISULFID 731 740 BY SIMILARITY.
 FT DISULFID 748 759 BY SIMILARITY.
 FT DISULFID 753 768 BY SIMILARITY.
 FT DISULFID 770 779 BY SIMILARITY.
 FT DISULFID 786 797 BY SIMILARITY.
 FT DISULFID 791 806 BY SIMILARITY.
 FT DISULFID 808 817 BY SIMILARITY.
 FT DISULFID 824 835 BY SIMILARITY.
 FT DISULFID 829 844 BY SIMILARITY.
 FT DISULFID 846 855 BY SIMILARITY.
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 382 382 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 559 559 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 745 745 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 960 960 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 991 991 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1064 1064 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1218 AA; 77739F8928BB793C CRC64;
 Query Match 11.4%; Score 86; DB 1; Length 1218;
 Best Local Similarity 27.0%; Pred. No. 0.81;
 Matches 33; Conservative 9; Mismatches 34; Indels 46; Gaps 9;
 QY 39 PE---GGTCEVIAHRC-----C-NKRIERSQTVKSCLPKGVAGTTRNR----- 81
 DB 325 PEGYSGPNCE-IAEHACLSDPCNRCGSKETSGFCECSPGWGTCTSNIDDCSPNNC 383
 QY 82 ---PSCVDASIVIG-----KWW-----CEMEPCLEGECEKTLDPN-----SG 115
 DB 384 SHGCTQD--LVNGFKVCVPPQWTGTCQLDANECAKPCVNARSCNLIASYCDCLPG 441
 QY 116 WM 117
 DB 442 WM 443
 RESULT 12
 JAG1_RAT STANDARD; PRT; 1219 AA.
 ID JAG1_RAT Q63722; P70640;
 AC 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Jagged 1 precursor (Jaggedl).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Sciatic nerve;
 RX MEDLINE=95211842; PubMed=7697721;
 RA Lindvall C.E., Shawler C.J., Boulter J., Weinmaster G.;
 RT "Jagged: a mammalian ligand that activates Notch1.";
 RL Cell 80:909-917(1995).
 CC -!- FUNCTION: Ligand for multiple Notch receptors and involved in the
 CC mediation of Notch signaling. May be involved in cell-fate
 CC decisions during hematopoiesis. Enhances fibroblast growth
 CC factor-induced angiogenesis (in vitro). Seems to be involved in
 CC early and late stages of mammalian cardiovascular development.
 CC Inhibits myoblast differentiation. May regulate fibroblast
 CC growth factor-induced angiogenesis.
 CC -!- SUBUNIT: Interacts with NOTCH1, NOTCH2 AND NOTCH3 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Widely expressed in a variety of tissues.

FT	DISULFID	324	333	BY SIMILARITY.
FT	DISULFID	340	351	BY SIMILARITY.
FT	DISULFID	345	360	BY SIMILARITY.
FT	DISULFID	362	371	BY SIMILARITY.
FT	DISULFID	378	389	BY SIMILARITY.
FT	DISULFID	383	398	BY SIMILARITY.
FT	DISULFID	400	409	BY SIMILARITY.
FT	DISULFID	416	427	BY SIMILARITY.
FT	DISULFID	421	436	BY SIMILARITY.
FT	DISULFID	438	447	BY SIMILARITY.
FT	DISULFID	454	464	BY SIMILARITY.
FT	DISULFID	458	473	BY SIMILARITY.
FT	DISULFID	475	484	BY SIMILARITY.
FT	DISULFID	491	502	BY SIMILARITY.
FT	DISULFID	496	511	BY SIMILARITY.
FT	DISULFID	513	522	BY SIMILARITY.
FT	DISULFID	529	540	BY SIMILARITY.
FT	DISULFID	534	549	BY SIMILARITY.
FT	DISULFID	551	560	BY SIMILARITY.
FT	DISULFID	578	605	BY SIMILARITY.
FT	DISULFID	599	615	BY SIMILARITY.
FT	DISULFID	617	626	BY SIMILARITY.
FT	DISULFID	633	644	BY SIMILARITY.
FT	DISULFID	638	653	BY SIMILARITY.
FT	DISULFID	655	664	BY SIMILARITY.
FT	DISULFID	671	682	BY SIMILARITY.
FT	DISULFID	676	691	BY SIMILARITY.
FT	DISULFID	693	702	BY SIMILARITY.
FT	DISULFID	709	720	BY SIMILARITY.
FT	DISULFID	714	729	BY SIMILARITY.
FT	DISULFID	731	740	BY SIMILARITY.
FT	DISULFID	748	759	BY SIMILARITY.
FT	DISULFID	753	768	BY SIMILARITY.
FT	DISULFID	770	779	BY SIMILARITY.
FT	DISULFID	786	797	BY SIMILARITY.
FT	DISULFID	791	806	BY SIMILARITY.
FT	DISULFID	808	817	BY SIMILARITY.
FT	DISULFID	824	835	BY SIMILARITY.
FT	DISULFID	829	844	BY SIMILARITY.
FT	DISULFID	846	855	BY SIMILARITY.
FT	CARBOHYD	143	143	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	217	217	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	382	382	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	559	559	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	745	745	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	960	960	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	991	991	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1045	1045	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1064	1064	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	SEQUENCE	1219 AA; 134325 MW; 65D4FC238AE204 CRC64;		

Query Match 11.4%; Score 86; DB 1; Length 1219;
Best Local Similarity 27.0%; Pred. No. 0.81;
Matches 33; Conservative 9; Mismatches 34; Indels 46; Gaps 9;

QY	39	PE---GGTCEVIAHRC-----C-NKNRIERSQTVKSCSLPGKVAGTRNR-----	81
Db	325	PEGYSGNCE-IAEHACLSDPCHNRSGKRETSGGFECECPGWTGPTCTSNIDDCSPNNC	383
QY	82	---PSCVDASIVTG-----KWW-----CEMEPCLEGECKTLFDN-----	115
Db	384	SHGTTQCD--LVNGFKVCVPQWTGKTQCLDANECAKPCVNARSCKNLIASYVCDCLPG	441
QY	116	WM 117	
Db	442	WM 443	

RESULT 13
FBPI_STRPU
ID FBPI_STRPU STANDARD; PRT; 1064 AA.
AC P10079;
DT 01-MAR-1989 (Rel. 10, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Fibropellin I precursor (Epidermal growth factor-related protein 1)
 DE (UEGF-1).
 GN EGF1.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinozoa; Echinacea; Echinoidea; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=90112459; PubMed=2514273;
 RX Delgadillo-Reynoso M.G., Rollo D.R., Hursh D.A., Raff R.A.;
 RA "Structural analysis of the UEGF gene in the sea urchin
 RT strongylocentrotus purpuratus reveals more similarity to vertebrate
 RT than to invertebrate genes with EGF-like repeats.";
 RL J. Mol. Evol. 29:314-327(1989).
 RN [2]
 RN SEQUENCE OF 279-476 AND 791-1064 FROM N.A.
 RP MEDLINE=87319677; PubMed=3498216;
 RX Hursh D.A., Andrews M.E., Raff R.A.;
 RA "A sea urchin gene encodes a polypeptide homologous to epidermal
 RT growth factor.";
 RL Science 237:1487-1490(1987).
 RN [3]
 RN AVIDIN-LIKE DOMAIN.
 RP MEDLINE=89196806; PubMed=2784773;
 RX Hunt L.T., Barker W.C.;
 RA "Avidin-like domain in an epidermal growth factor homolog from a sea
 RT urchin.";
 RL FASEB J. 3:1760-1764(1989).
 RN [4]
 RN CHARACTERIZATION.
 RP MEDLINE=91285254; PubMed=2060714;
 RX Bisgrove B.W., Andrews M.E., Raff R.A.;
 RA "Fibropellins, products of an EGF repeat-containing gene, form a
 RT unique extracellular matrix structure that surrounds the sea urchin
 RT embryo.";
 RL Dev. Biol. 146:89-99(1991).
 CC -|- FUNCTION: Form the apical lamina, a component of the extracellular
 CC matrix.
 CC -|- SUBCELLULAR LOCATION: EXTRACELLULAR, IN VESICLES IN THE CYTOPLASM
 CC OF UNFERTILIZED EGGS, THEN TO THE BASE OF THE HYALIN LAYER
 CC THROUGHOUT DEVELOPMENT AND FINALLY IN THE APICAL LAMINA IN LATE
 CC EMBRYOS AND EARLY LARVAE.
 CC -|- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=IA;
 CC IsoId=PI0079-1; Sequence=Displayed;
 CC Name=IB;
 CC IsoId=PI0079-2; Sequence=VSP_000451;
 CC -|- DEVELOPMENTAL STAGE: Moderate levels in unfertilized eggs and
 CC during early cleavage, then rapidly increases in abundance between
 CC late morula and mesenchyme blastula stages to maximal levels
 CC maintained through subsequent stages. Expressed both maternally
 CC and zygotically.
 CC -|- SIMILARITY: Contains 21 EGF-like domains.
 CC -|- SIMILARITY: Contains 1 CUB domain.
 CC -|- SIMILARITY: THE C-TERMINAL DOMAIN OF THIS PROTEIN IS SIMILAR TO
 CC AVIDIN/SREPTAVIDIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L08692; AAA62164.1; -
 CC EMBL; L08692; AAA62163.1; -
 CC EMBL; X17530; CAA35571.1; -

DR EMBL; M17421; AAA30050.1; -
 DR EMBL; X17533; CAA35573.1; -
 DR PIR; A40136; A40136.
 DR HSSP; P01132; 1EGF.
 DR InterPro; IPR000152; Asx_hydroxyl_s.
 DR InterPro; IPR005469; Avidin.
 DR InterPro; IPR005468; Avidin/str.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001438; EGF_II.
 DR InterPro; IPR006209; EGF_Like.
 DR Pfam; PF01382; Avidin; 1.
 DR Pfam; PF00431; CUB; 1.
 DR Pfam; PF00008; EGF; 21.
 DR PRINTS; PR00709; AVIDIN.
 DR PRINTS; PR00010; EGFBL00D.
 DR SMART; SM00042; CUB; 1.
 DR SMART; SM00179; EGF_Ca; 20.
 DR PROSITE; PS00010; ASX_HYDROXYL; 19.
 DR PROSITE; PS00577; AVIDIN; 1.
 DR PROSITE; PS01180; CUB; 1.
 DR PROSITE; PS00022; EGF_1; 19.
 DR PROSITE; PS01186; EGF_2; 19.
 DR PROSITE; PS50036; EGF_3; 21.
 DR PROSITE; PS01187; EGF_Ca; 18.
 DR Biotin; Alternative splicing; EGF-like domain; Repeat; Signal;
 KW Glycoprotein; Calcium-binding.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 1064 FIBROPELLIN 1.
 FT DOMAIN 20 55 EGF-LIKE 1.
 FT DOMAIN 62 175 CUB.
 FT DOMAIN 176 212 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 214 250 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 252 288 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 290 326 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 328 364 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 366 402 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 404 440 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 442 478 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 480 516 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 518 554 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 556 592 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 594 630 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 632 668 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 670 706 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 708 744 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 746 782 EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 784 820 EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 822 858 EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 860 896 EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 898 934 EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 936 1064 AVIDIN-LIKE.
 FT DOMAIN 23 34 BY SIMILARITY.
 FT DISULFID 28 43 BY SIMILARITY.
 FT DISULFID 45 54 BY SIMILARITY.
 FT DISULFID 180 191 BY SIMILARITY.
 FT DISULFID 185 200 BY SIMILARITY.
 FT DISULFID 202 211 BY SIMILARITY.
 FT DISULFID 218 229 BY SIMILARITY.
 FT DISULFID 223 238 BY SIMILARITY.
 FT DISULFID 240 249 BY SIMILARITY.
 FT DISULFID 256 267 BY SIMILARITY.
 FT DISULFID 261 276 BY SIMILARITY.
 FT DISULFID 278 287 BY SIMILARITY.
 FT DISULFID 294 305 BY SIMILARITY.
 FT DISULFID 299 314 BY SIMILARITY.
 FT DISULFID 316 325 BY SIMILARITY.
 FT DISULFID 332 343 BY SIMILARITY.
 FT DISULFID 337 352 BY SIMILARITY.
 FT DISULFID 354 363 BY SIMILARITY.
 FT DISULFID 370 381 BY SIMILARITY.
 FT DISULFID 375 390 BY SIMILARITY.

[illegible]

from the membrane (By similarity).
 -!- PTM: Phosphorylated (By similarity).
 -!- POLYMORPHISM: The poly-Leu region of NOTCH4 (in the signal peptide) is polymorphic and the number of Leu varies in the population (from 6 to 12).
 -!- SIMILARITY: Belongs to the NOTCH family.
 -!- SIMILARITY: Contains 28 EGF-like domains.
 -!- SIMILARITY: Contains 3 lin/Notch repeats.
 -!- SIMILARITY: Contains 5 ANK repeats.
 -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts in position 1438 to 1463.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

 EMBL; D63395; BAA09708.1; ALT_FRAME.
 EMBL; D86566; BAA13116.1; -.
 EMBL; U95299; AAC32288.1; -.
 EMBL; U89335; AAC63097.1; -.
 EMBL; AB023961; BAB20317.1; -.
 EMBL; AB024520; BAA88951.1; -.
 EMBL; AB024578; BAA88952.1; -.
 HSSP; P08709; 1BF9.
 Genew; HGNC:7884; NOTCH4.
 MIM; 164951; -.
 InterPro; IPR002110; ANK.
 InterPro; IPR000152; ASX_hydroxyl_S.
 InterPro; IPR000742; EGF_2.
 InterPro; IPR001881; EGF_Ca.
 InterPro; IPR001438; EGF_II.
 InterPro; IPR006209; EGF_like.
 InterPro; IPR002049; Laminin_EGF.
 InterPro; IPR008297; Notch.
 InterPro; IPR000800; Notch_dom.
 Pfam; PF00023; ank; 6.
 Pfam; PF00066; notch; 2.
 Pfam; PF00008; EGF; 26.
 PIRSF; PIRSF02279; Notch; 1.
 PRINTS; PR00010; EGFLOOD.
 PRINTS; PR00011; EGF_LAMININ.
 PRINTS; PR01452; NOTCH.
 SMART; SM00248; ANK; 5.
 SMART; SM00179; EGF_CA; 11.
 SMART; SM00004; NL; 3.
 PROSITE; PS0297; ANK_REPEAT_REGION; 1.
 PROSITE; PS0088; ANK_REPEAT; 5.
 PROSITE; PS00010; ASX_HYDROXYL; 11.
 PROSITE; PS00022; EGF_1; 28.
 PROSITE; PS01186; EGF_2; 21.
 PROSITE; PS00026; EGF_3; 28.
 PROSITE; PS01187; EGF_CA; 9.
 Receptor; transcription regulation; Activator; Differentiation;
 Developmental protein; Repeat; ANK repeat; EGF-like domain;
 Transmembrane; Glycoprotein; Signal; Phosphorylation; Polymorphism;
 Triplet repeat expansion; Alternative splicing.
 SIGNAL 1 23 POTENTIAL.
 CHAIN 24 2003 NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG 4.
 CHAIN 1432 2003 NOTCH EXTRACELLULAR TRUNCATION
 (BY SIMILARITY).
 CHAIN 1467 2003 NOTCH INTRACELLULAR DOMAIN
 (BY SIMILARITY).
 DOMAIN 24 1447 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 1448 1468 POTENTIAL.
 DOMAIN 1469 2003 CYTOPLASMIC (POTENTIAL).
 DOMAIN 24 63 EGF-LIKE 1.
 DOMAIN 64 115 EGF-LIKE 2.
 DOMAIN 118 155 EGF-LIKE 3.
 DOMAIN 156 192 EGF-LIKE 4.

FT	DOMAIN	194	232	EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	234	274	EGF-LIKE 6.
FT	DOMAIN	276	312	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	314	353	EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	355	391	EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	392	430	EGF-LIKE 10.
FT	DOMAIN	432	473	EGF-LIKE 11.
FT	DOMAIN	475	511	EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	513	549	EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	551	587	EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	589	625	EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	626	659	EGF-LIKE 16.
FT	DOMAIN	661	689	EGF-LIKE 17.
FT	DOMAIN	691	727	EGF-LIKE 18.
FT	DOMAIN	729	765	EGF-LIKE 19.
FT	DOMAIN	767	803	EGF-LIKE 20.
FT	DOMAIN	806	842	EGF-LIKE 21.
FT	DOMAIN	844	880	EGF-LIKE 22.
FT	DOMAIN	882	928	EGF-LIKE 23.
FT	DOMAIN	930	966	EGF-LIKE 24.
FT	DOMAIN	968	1004	EGF-LIKE 25.
FT	DOMAIN	1006	1044	EGF-LIKE 26.
FT	DOMAIN	1046	1085	EGF-LIKE 27.
FT	DOMAIN	1087	1126	EGF-LIKE 28.
FT	DOMAIN	1130	1171	EGF-LIKE 29.
FT	DOMAIN	1172	1476	POLY-ARG.
FT	REPEAT	1165	1212	LIN/NOTCH 1.
FT	REPEAT	1213	1246	LIN/NOTCH 2.
FT	REPEAT	1247	1286	LIN/NOTCH 3.
FT	REPEAT	1633	1665	ANK 1.
FT	REPEAT	1666	1698	ANK 2.
FT	REPEAT	1700	1732	ANK 3.
FT	REPEAT	1733	1765	ANK 4.
FT	REPEAT	1766	1798	ANK 5.
FT	DISULFID	28	41	BY SIMILARITY.
FT	DISULFID	35	51	BY SIMILARITY.
FT	DISULFID	53	62	BY SIMILARITY.
FT	DISULFID	68	80	BY SIMILARITY.
FT	DISULFID	74	103	BY SIMILARITY.
FT	DISULFID	105	114	BY SIMILARITY.
FT	DISULFID	122	133	BY SIMILARITY.
FT	DISULFID	127	143	BY SIMILARITY.
FT	DISULFID	145	154	BY SIMILARITY.
FT	DISULFID	160	171	BY SIMILARITY.
FT	DISULFID	165	180	BY SIMILARITY.
FT	DISULFID	182	191	BY SIMILARITY.
FT	DISULFID	198	211	BY SIMILARITY.

Query Match 11.2%; Score 84.5; DB 1; Length 2003;
 Best Local Similarity 31.0%; Pred. No. 1.9;
 Matches 39; Conservative 6; Mismatches 42; Indels 39; Gaps 10;

QY	18	CAMLLCH-GS----	LOHTFQHHLRP--	EGGTCEV----	IAAHRCCNKNRIEERSQTVK 66
Db	479	CLSPCHPGSTCLDLATP--	HCLCPGLEQLCEVETNECASAPCLNHADCHDLINGFQ 536		
QY	67	CSCLPKVGAGTRNRPS-	VDASIVIGKWCNEPCLGEE-----	CKILPDNS 114	
Db	537	CICLPG-FSGT-----	RCEDIDE-----	CRSSPCANGGQQCDQPGAFHCKCLPGFE 582	
QY	115	GNMCA 120			
Db	583	GPFCQT 588			

RESULT 15
 JAG2_MOUSE STANDARD; PRT; 1247 AA.
 ID JAG2_MOUSE
 AC Q9Y8S; O55139; O70219;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Jagged 2 precursor (Jagged2).

GN JAG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss Webster / NIH;
RA Tsai S.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE OF 302-819 FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=98051918; PubMed=9341252;
RX Lan F., Jiang R., Shawber C., Weinmaster G., Gridley T.;
RA "The Jagged2 gene maps to chromosome 12 and is a candidate for the lgl
RT and sm mutations.";
RL Mamm. Genome 8:875-876(1997).
[3]
RN SEQUENCE OF 325-759 FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=98145947; PubMed=9486542;
RX Valsecchi C., Ghezzi C., Ballabio A., Rugarli E.I.;
RA "JAGGED2: a putative Notch ligand expressed in the apical ectodermal
RT ridge and in sites of epithelial-mesenchymal interactions.";
RL Mech. Dev. 69:203-207(1997).
[4]
RN TISSUE SPECIFICITY.
RP MEDLINE=97459705; PubMed=9315665;
RX Luo B., Aster J.C., Hasserjian R.P., Kuo F., Sklar J.;
RA "Isolation and functional analysis of a cDNA for human Jagged2, a gene
RT encoding a ligand for the Notch1 receptor.";
RL Mol. Cell. Biol. 17:6057-6067(1997).
CC -!- FUNCTION: Putative Notch ligand involved in the mediation of Notch
CC signaling. Plays an essential role during limb, craniofacial and
CC thymic development. May be involved in myogenesis and in the
CC development of peripheral and central nervous systems.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Found to be highest in fetal thymus,
CC epidermis, foregut dorsal root ganglia and inner ear. In 2-week-
CC old mice, abundant in heart, lung, thymus, skeletal muscle, brain
CC and testis. Expression overlaps partially with Notch1 expression.
CC -!- DEVELOPMENTAL STAGE: At 13 dpc, found in paravertebral vessels and
CC dorsal root ganglia. At 14 dpc, in oropharyngeal epithelium,
CC developing thymus and in the muscles of the tongue. By 15 dpc, in
CC many tissues.
CC -!- SIMILARITY: Contains 16 EGF-like domains.
CC -!- SIMILARITY: Contains 1 DSL domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF038572; AAF16411.1; -;
CC EMBL; AF010137; AAC14010.1; -;
CC EMBL; Y14495; CAA74835.1; -;
CC HSSP; P00743; 1CCF.
CC MGD; MGI:1098270; Jag2.
CC GO; GO:0005887; C:integral to plasma membrane; IDA.
CC GO; GO:0008083; F:growth factor activity; ISS.
CC GO; GO:0005112; F:Notch binding; IPI.
CC GO; GO:0007049; P:cell cycle; ISS.
CC GO; GO:0001544; P:cell differentiation; NAS.
CC GO; GO:0001709; P:cell fate determination; NAS.
CC GO; GO:0007267; P:cell-cell signaling; NAS.
CC GO; GO:0042492; P:gamma-delta T-cell differentiation; IMP.
CC GO; GO:0009912; P:hair cell fate commitment; IMP.
CC GO; GO:0007605; P:hearing; NAS.
CC GO; GO:0030326; P:limb morphogenesis; ISS.
CC -----
DR GO; GO:0007219; P:N signaling pathway; ISS.
DR GO; GO:0030334; P:regulation of cell migration; ISS.
DR GO; GO:0042127; P:regulation of cell proliferation; ISS.
DR GO; GO:0007283; P:spermatogenesis; ISS.
DR GO; GO:0045061; P:thymic T-cell selection; ISS.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR001774; DSL.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01414; DSL; 1.
DR Pfam; PF00008; EGF; 14.
DR PRINTS; PR00010; EGFBL00D.
DR SMART; SM00051; DSL; 1.
DR SMART; SM00179; EGF_CA; 9.
DR SMART; SM00214; VMC; 1.
DR PROSITE; PS000152; ASX_HYDROXYL; 10.
DR PROSITE; PS00022; EGF_1; 16.
DR PROSITE; PS01186; EGF_2; 11.
DR PROSITE; PS00026; EGF_3; 15.
DR PROSITE; PS01187; EGF_CA; 7.
KW Calcium-binding; EGF-like domain; Glycoprotein; Developmental protein;
KW Repeat; Signal; Transmembrane.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 1247 JAGGED 2.
FT DOMAIN 24 1084 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1085 1107 POTENTIAL.
FT DOMAIN 1108 1247 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 178 240 DSL.
FT DOMAIN 241 274 EGF-LIKE 1.
FT DOMAIN 275 305 EGF-LIKE 2.
FT DOMAIN 307 345 EGF-LIKE 3.
FT DOMAIN 347 383 EGF-LIKE 4.
FT DOMAIN 385 421 EGF-LIKE 5.
FT DOMAIN 423 459 EGF-LIKE 6.
FT DOMAIN 461 496 EGF-LIKE 7.
FT DOMAIN 498 534 EGF-LIKE 8.
FT DOMAIN 536 572 EGF-LIKE 9.
FT DOMAIN 574 634 EGF-LIKE 10 (ATYPICAL).
FT DOMAIN 636 672 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 674 710 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 712 748 EGF-LIKE 13.
FT DOMAIN 751 787 EGF-LIKE 14.
FT DOMAIN 789 825 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 827 863 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 245 256 BY SIMILARITY.
FT DISULFID 249 262 BY SIMILARITY.
FT DISULFID 264 273 BY SIMILARITY.
FT DISULFID 276 287 BY SIMILARITY.
FT DISULFID 282 293 BY SIMILARITY.
FT DISULFID 295 304 BY SIMILARITY.
FT DISULFID 311 323 BY SIMILARITY.
FT DISULFID 317 333 BY SIMILARITY.
FT DISULFID 335 344 BY SIMILARITY.
FT DISULFID 351 362 BY SIMILARITY.
FT DISULFID 356 371 BY SIMILARITY.
FT DISULFID 373 382 BY SIMILARITY.
FT DISULFID 389 400 BY SIMILARITY.
FT DISULFID 394 409 BY SIMILARITY.
FT DISULFID 411 420 BY SIMILARITY.
FT DISULFID 427 438 BY SIMILARITY.
FT DISULFID 432 447 BY SIMILARITY.
FT DISULFID 449 458 BY SIMILARITY.
FT DISULFID 465 475 BY SIMILARITY.
FT DISULFID 469 484 BY SIMILARITY.
FT DISULFID 486 495 BY SIMILARITY.
FT DISULFID 502 513 BY SIMILARITY.
FT DISULFID 507 522 BY SIMILARITY.
FT DISULFID 524 533 BY SIMILARITY.
FT DISULFID 540 551 BY SIMILARITY.
FT DISULFID 545 560 BY SIMILARITY.

FT	DISULFID	562	571	BY SIMILARITY.
FT	DISULFID	589	612	POTENTIAL.
FT	DISULFID	606	622	POTENTIAL.
FT	DISULFID	624	633	BY SIMILARITY.
FT	DISULFID	640	651	BY SIMILARITY.
FT	DISULFID	645	660	BY SIMILARITY.
FT	DISULFID	662	671	BY SIMILARITY.
FT	DISULFID	678	689	BY SIMILARITY.
FT	DISULFID	683	698	BY SIMILARITY.
FT	DISULFID	700	709	BY SIMILARITY.
FT	DISULFID	716	727	BY SIMILARITY.
FT	DISULFID	721	736	BY SIMILARITY.
FT	DISULFID	738	747	BY SIMILARITY.
FT	DISULFID	755	766	BY SIMILARITY.
FT	DISULFID	760	775	BY SIMILARITY.
FT	DISULFID	777	786	BY SIMILARITY.
FT	DISULFID	793	804	BY SIMILARITY.
FT	DISULFID	798	813	BY SIMILARITY.
FT	DISULFID	815	824	BY SIMILARITY.
FT	DISULFID	831	842	BY SIMILARITY.
FT	DISULFID	836	851	BY SIMILARITY.
FT	DISULFID	853	862	BY SIMILARITY.
FT	CARBOHYD	153	153	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	570	570	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	619	619	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	752	752	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1060	1060	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	302	302	L -> M (IN REF. 2).
FT	CONFLICT	461	461	N -> T (IN REF. 2).
FT	CONFLICT	469	478	COHGGTCKDL -> VSAWHLQGP (IN REF. 2).
FT	CONFLICT	492	492	G -> V (IN REF. 2).
FT	CONFLICT	546	546	L -> F (IN REF. 2).
FT	CONFLICT	549	549	A -> V (IN REF. 2).
FT	CONFLICT	735	738	RCAC -> PAR (IN REF. 3).
FT	CONFLICT	809	809	N -> H (IN REF. 2).
FT	CONFLICT	812	812	R -> A (IN REF. 2).
SQ	SEQUENCE	1247	AA; 134726 MW; 1D80C8628FAFEEC CRC64;	

Query Match 11.1%; Score 84; DB 1; Length 1247;
 Best Local Similarity 24.6%; Pred. No. 1.3;
 Matches 30; Conservative 12; Mismatches 30; Indels 50; Gaps 9;

QY	23	CHGSLQHTFQOHLHREPGTTC-EVIAAHRC-----CNKNRI- 58
Db	465	CHGQCQH-----GGTCKDLVNGYQCVPRGFGRRHCELEYDKCASSPCRRGGIC 513
QY	59	EERSQTVKCSCLPGKVAGTTNRPS-CVDASIVICKWKCMEPCGECECKTLPDNGWM 117
Db	514	EDLVDGFRCHC-PRGLSGL-----HCEVDMDL-----CEPSPCLNGARCYNL--EGDYY 559
QY	118	CA 119
Db	560	CA 561

Search completed: April 8, 2004, 13:16:40
 Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 13:15:04 ; Search time 22 Seconds
(without alignments)
312.102 Million cell updates/sec

Title: US-09-763-335-2
Perfect score: 754
Sequence: 1 MAMVSAMSWVLYLWISACAM.....SGMWCATGNKIKTRIHPT 133

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/ptotus-COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	19.9	29	3	US-08-905-223-451
2	110.5	14.7	1248	3	US-08-892-046-6
3	96	12.7	1964	4	US-09-467-997-1
4	91.5	12.1	233	4	US-09-216-393B-110
5	86.5	11.5	156	4	US-09-252-991A-22898
6	86.5	11.5	1139	1	US-08-537-210A-4
7	86.5	11.5	1139	3	US-09-113-825-4
8	86.5	11.5	2703	1	US-08-185-432-19
9	86.5	11.5	2703	4	US-08-899-232-4
10	86	11.4	500	4	US-09-423-753-2
11	86	11.4	659	4	US-09-423-753-3
12	86	11.4	685	3	US-08-872-855-2
13	86	11.4	685	4	US-09-423-753-25
14	86	11.4	1010	3	US-08-882-046-7
15	86	11.4	1036	4	US-09-068-740A-6
16	86	11.4	1187	4	US-09-068-740A-7
17	86	11.4	1208	4	US-09-199-865-1
18	86	11.4	1218	2	US-08-400-159-6
19	86	11.4	1218	3	US-08-611-729A-6
20	86	11.4	1218	3	US-08-882-046-2
21	86	11.4	1218	3	US-09-214-278-7
22	86	11.4	1218	4	US-09-068-740A-11
23	86	11.4	1218	4	US-09-855-722-7
24	86	11.4	1219	3	US-08-882-046-5
25	84	11.1	2523	1	US-08-185-432-18
26	84	11.1	2523	4	US-08-899-232-3
27	82.5	10.9	2471	1	US-08-185-432-16

28	82.5	10.9	2471	1	US-08-083-590A-19
29	82.5	10.9	2471	3	US-08-532-384-19
30	82.5	10.9	2471	4	US-08-899-232-1
31	82	10.9	278	4	US-09-724-864-52
32	82	10.9	4544	1	US-08-469-486-52
33	82	10.9	4544	2	US-08-469-658-52
34	81.5	10.8	1148	3	US-08-882-046-4
35	81.5	10.8	2556	1	US-08-083-590A-20
36	81.5	10.8	2556	3	US-08-532-384-20
37	80.5	10.7	2321	4	US-09-230-652-2
38	80	10.6	2556	1	US-08-185-432-17
39	80	10.6	2556	4	US-08-899-232-2
40	79	10.5	1193	2	US-08-400-159-10
41	79	10.5	1193	3	US-08-611-729A-10
42	77.5	10.3	520	4	US-09-068-740A-3
43	77.5	10.3	702	4	US-09-068-740A-4
44	77.5	10.3	723	4	US-09-068-740A-9
45	77.5	10.3	723	4	US-09-423-753-27

ALIGNMENTS

RESULT 1
US-08-905-223-451
; Sequence 451, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 451:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -25...-1
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 5.8
; OTHER INFORMATION: seq LWSACMLLCHG/SL
US-08-905-223-451

Query Match 19.9%; Score 150; DB 3; Length 29; Indels 0; Gaps 0;
Best Local Similarity 100.0%; Pred. No. 1.9e-09; Mismatches 0; Indels 0; Gaps 0;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAMVSAMSWLYLWISACAMLLCHGSLQ 28
DB 1 MAMVSAMSWLYLWISACAMLLCHGSLQ 28

RESULT 2
US-08-882-046-6
; Sequence 6, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1248 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-882-046-6

Query Match 14.7%; Score 110.5; DB 3; Length 1248;
Best Local Similarity 27.0%; Pred. No. 0.0022;
Matches 33; Conservative 13; Mismatches 27; Indels 49; Gaps 9;
QY 23 CHGSLQHTFOQHHLHRRPGGTC-EVIAAHR-----CNKNRI- 58
DB 465 CHGQCH-----GGTKDLVNGYQVCPRGFGHCELEYKCASSPCRRGGIC 513

QY 59 EERSQTVKSCLPKGVAGTTRNRPS-CVDASIVIGKWWCEMEPCLEGECKTLPDNSGWM 117
DB 514 EBLVDGFRCHC-FRGLSG-----PLCEVDVDL-----WCEPNPCLNGARCYNLEDD---YY 560

QY 118 CA 119
DB 561 CA 562

RESULT 3
US-09-467-997-1
; Sequence 1, Application US/09467997
; Patent No. 6379925

GENERAL INFORMATION:
; APPLICANT: Kitajewski, Jan
; APPLICANT: Uyttendaele, Hendrik
; TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
; FILE REFERENCE: 53863-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/467,997
; CURRENT FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1964
; TYPE: PRT
; ORGANISM: mouse
US-09-467-997-1

Query Match 12.7%; Score 96; DB 4; Length 1964;
Best Local Similarity 26.8%; Pred. No. 0.13;
Matches 33; Conservative 8; Mismatches 48; Indels 34; Gaps 6;
QY 15 ISACAMLLC-HGSLQHTFOQHHLHRRPGGTC-----EVIAAHR--CCNKNRI 58
DB 689 LGGCISTPCAAGGTGCHP-----QPSGYNCTCPAGYMGTCSEEVTAACHSGPCLNGGSC 741

QY 59 EERSQTVKSCLPKGVAGTTRNRPS-CVDASIVIGKWWCEMEPCLEGECKTLPDNSGWM 118
DB 742 SIRPEGYSTCLP---SHTGRHCQTAVD-----HCVSASCLNGGTCVNKEGTFFCLC 790

QY 119 ATG 121
DB 791 ATG 793

RESULT 4
US-09-216-393B-110
; Sequence 110, Application US/09216393B
; Patent No. 6514694
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 110
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
US-09-216-393B-110

Query Match 12.1%; Score 91.5; DB 4; Length 233;
Best Local Similarity 25.9%; Pred. No. 0.036;
Matches 28; Conservative 6; Mismatches 39; Indels 35; Gaps 4;
QY 18 CAMLLCHGSLQHTFOQHHLHRRPGGTC-EVIAAHR-----CNKN-----RIERS 62
DB 78 CQDCACH-----GKTCETKHCCHCCINDSCNGHGTCTNTNTNCAGF 122

QY 63 QTVKSCLPKGVAGTTRNRPS-CVDASIVIGKWWCEMEPCLEGECKTLPDNSGWM 110
DB 123 AGTNCSSSEKCSGKT-----CLSGHCNPATGACVCDPCHTGERCETL 165

RESULT 5
US-09-252-991A-22898
; Sequence 22898, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22898
LENGTH: 156
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22898

Query Match      11.5%; Score 86.5; DB 4; Length 156;
Best Local Similarity 25.0%; Pred. No. 0.078;
Matches 32; Conservative 9; Mismatches 64; Indels 23; Gaps 5;

QY 6 AMSWVLYLWISACAMLLCHGSLQHTQOHLHREPGTCEVIAAHRCCNKNRIEERSQTV 65
DB 12 AASW--KSWSSACS--TSTGCLPTYAPARRPSPARGSSSKAVAAARCCNATTPCSNSRSP 67
QY 66 K-----GSCLPKGVAGTTR-----NRPSQVDASIVIGKWCMEPCLEGECKT 109
DB 68 NPCCLCWSSGTCRCCLPISTDRMTWPASATRESTPSAPV---RWPRPPACTSMRCSR 124
QY 110 LPNSGWM 117
DB 125 RSARRGWI 132

RESULT 6
US-08-537-210A-4
; Sequence 4, Application US/08537210A
; Patent No. 5780300
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Fortini, Mark
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: MANIPULATION OF NON-TERMINALLY
; TITLE OF INVENTION: DIFFERENTIATED CELLS USING THE NOTCH PATHWAY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,210A
; FILING DATE: 29-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-027
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1139 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein

Query Match      11.5%; Score 86.5; DB 3; Length 1139;
```

```
FEATURE:
; NAME/KEY: Dros N
; LOCATION: 1189...2327
; OTHER INFORMATION: Highly conserved ankyrin repeat
; OTHER INFORMATION: region of No. 5780300ch
US-08-537-210A-4

Query Match      11.5%; Score 86.5; DB 1; Length 1139;
Best Local Similarity 25.9%; Pred. No. 0.74;
Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;

QY 40 EGGTCEV-----IAAHRCCNKNRIEERSQTVKCSCLPGKVA-----GTTNRPS 83
DB 26 QGNCNELNIDDCAPNQCQGGTCHDRVNMNFSCPPGTWGIICEINKDKCKPGACHNNGS 85
QY 84 CYDASIVIGKWCMEPCLEGECK 108
DB 86 CIDR---VGFECVQCPGFVGARCE 107

RESULT 7
US-09-113-825-4
; Sequence 4, Application US/09113825
; Patent No. 6149902
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Fortini, Mark
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: MANIPULATION OF NON-TERMINALLY
; TITLE OF INVENTION: DIFFERENTIATED CELLS USING THE NOTCH PATHWAY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/113,825
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/537,210
; FILING DATE: 29-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-027
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1139 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Dros N
; LOCATION: 1189...2327
; OTHER INFORMATION: Highly conserved ankyrin repeat
; OTHER INFORMATION: region of No. 6149902ch
US-09-113-825-4

Query Match      11.5%; Score 86.5; DB 3; Length 1139;
```

Best Local Similarity 25.9%; Pred. No. 0.74;
Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;

QY 40 EGGTCV-----IAHRCCKNRIEERSQTVKCSCLPGKVA-----GTTNRPS 83
Db 26 QGNCNELNDDCAPNCPQGGTCHDRVMNFSCSPGTMGIICEINKDDCKPGACHNNGS 85

QY 84 CVDASIVIGKWCMEPCLEGECK 108
Db 86 CIDR---VGGFECVCQGFVGARCE 107

RESULT 8
US-08-185-432-19
; Sequence 19, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/185,432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2703 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-185-432-19

Query Match 11.5%; Score 86.5; DB 1; Length 2703;
Best Local Similarity 25.9%; Pred. No. 2;
Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;

QY 40 EGGTCV-----IAHRCCKNRIEERSQTVKCSCLPGKVA-----GTTNRPS 83
Db 1214 QGNCNELNDDCAPNCPQGGTCHDRVMNFSCSPGTMGIICEINKDDCKPGACHNNGS 1273

QY 84 CVDASIVIGKWCMEPCLEGECK 108
Db 1274 CIDR---VGGFECVCQGFVGARCE 1295

RESULT 9
US-08-899-232-4
; Sequence 4, Application US/08899232

Best Local Similarity 25.9%; Pred. No. 0.74;
Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;

QY 40 EGGTCV-----IAHRCCKNRIEERSQTVKCSCLPGKVA-----GTTNRPS 83
Db 26 QGNCNELNDDCAPNCPQGGTCHDRVMNFSCSPGTMGIICEINKDDCKPGACHNNGS 85

QY 84 CVDASIVIGKWCMEPCLEGECK 108
Db 86 CIDR---VGGFECVCQGFVGARCE 107

RESULT 11
US-09-423-753-3
; Sequence 3, Application US/09423753
; Patent No. 6664098
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; TITLE OF INVENTION: NOVEL DIFFERENTIATION INHIBITOR
; FILE REFERENCE: KP-8693
; CURRENT APPLICATION NUMBER: US/09/423,753

Query Match 11.4%; Score 86; DB 4; Length 500;
Best Local Similarity 25.0%; Pred. No. 0.33;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

QY 14 WISAC-----AMLLCHSLQHTFQHHLHREPGTCEVIAHRCCKNRIEERSQTVKCS 69
Db 242 WQCTDEGHWGLFCDQDLNYC---THSPCKNGATC-----SNGQSRYSY--CTC 286

QY 70 LPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATG 121
Db 287 RPYGTG-----VDCLELSE--CDSNPCRNCGSSCKQEDGYHCLCPGP 327

RESULT 10
US-09-423-753-2
; Sequence 2, Application US/09423753
; Patent No. 6664098
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; TITLE OF INVENTION: NOVEL DIFFERENTIATION INHIBITOR
; FILE REFERENCE: KP-8693
; CURRENT APPLICATION NUMBER: US/09/423,753
; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: PCT/JP98/02104
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: JP 9/124064
; PRIOR FILING DATE: 1997-05-14
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-753-2

Query Match 11.5%; Score 86.5; DB 4; Length 2703;
Best Local Similarity 25.9%; Pred. No. 2;
Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;

QY 40 EGGTCV-----IAHRCCKNRIEERSQTVKCSCLPGKVA-----GTTNRPS 83
Db 1214 QGNCNELNDDCAPNCPQGGTCHDRVMNFSCSPGTMGIICEINKDDCKPGACHNNGS 1273

QY 84 CVDASIVIGKWCMEPCLEGECK 108
Db 1274 CIDR---VGGFECVCQGFVGARCE 1295

RESULT 11
US-09-423-753-3
; Sequence 3, Application US/09423753
; Patent No. 6664098
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; TITLE OF INVENTION: NOVEL DIFFERENTIATION INHIBITOR
; FILE REFERENCE: KP-8693
; CURRENT APPLICATION NUMBER: US/09/423,753

```
; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: PCT/JP98/02104
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: JP 9/124064
; PRIOR FILING DATE: 1997-05-14
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-753-3

Query Match 11.4%; Score 86; DB 4; Length 659;
Best Local Similarity 25.0%; Pred. No. 0.45;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

QY 14 WISAC-----AMLLCHGSLQHTFOQHLHRRPGGTCEVIAAHRCCNKNRIEERSQTVKSC 69
Db 242 WQCTDEGWLFCDDQDLYNC--THSPCKNGATC-----SNSGQSYT--CTC 286

QY 70 LPKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATG 121
Db 287 RPYGTG-----VDCELELSE--CDSNPCRNGSGCKQEDGYHCLCPPG 327

RESULT 12
US-08-872-855-2
; Sequence 2, Application US/08872855
; Patent No. 6121045
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean
; APPLICANT: Gearing, David
; TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/872,855
; FILING DATE: 11-JUN-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MAA-003.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 685 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-872-855-2

Query Match 11.4%; Score 86; DB 3; Length 685;
Best Local Similarity 25.0%; Pred. No. 0.47;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

QY 14 WISAC-----AMLLCHGSLQHTFOQHLHRRPGGTCEVIAAHRCCNKNRIEERSQTVKSC 69
```

```
Db 268 WQCTDEGWLFCDDQDLYNC--THSPCKNGATC-----SNSGQSYT--CTC 312
QY 70 LPKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATG 121
Db 313 RPYGTG-----VDCELELSE--CDSNPCRNGSGCKQEDGYHCLCPPG 353

RESULT 13
US-09-423-753-25
; Sequence 25, Application US/09423753
; Patent No. 6664098
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; TITLE OF INVENTION: NOVEL DIFFERENTIATION INHIBITOR
; FILE REFERENCE: KP-8693
; CURRENT APPLICATION NUMBER: US/09/423,753
; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: PCT/JP98/02104
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: JP 9/124064
; PRIOR FILING DATE: 1997-05-14
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-753-25

Query Match 11.4%; Score 86; DB 4; Length 685;
Best Local Similarity 25.0%; Pred. No. 0.47;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

QY 14 WISAC-----AMLLCHGSLQHTFOQHLHRRPGGTCEVIAAHRCCNKNRIEERSQTVKSC 69
Db 268 WQCTDEGWLFCDDQDLYNC--THSPCKNGATC-----SNSGQSYT--CTC 312

QY 70 LPKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATG 121
Db 313 RPYGTG-----VDCELELSE--CDSNPCRNGSGCKQEDGYHCLCPPG 353

RESULT 14
US-08-882-046-7
; Sequence 7, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
```

REFERENCE/DOCKET NUMBER: P-UW 2637
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1010 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-882-046-7

Query Match 11.4%; Score 86; DB 3; Length 1010;
Best Local Similarity 23.8%; Pred. No. 0.73; 34; Indels 64; Gaps 9;
Matches 34; Conservative 11; Mismatches 11; Mismatches 34; Indels 64; Gaps 9;
QY 37 HRP-EGGTCE-----VIAAHR-----C-NKNRIEERSQTVKSC 69
Db 303 HQPLNGGTGTCNTGPKYQSCPEGYSGPNCIAEHAACLSDPCHNRGCKETSLSGFECEC 362
QY 70 LPKGVAGTTNR-----PSCVDASIVIG-----KWW-----CEMEP 100
Db 363 SPGWGTGTCSTNIDDCSPNCSHGTCQD--LVNGFKVCVPPQWTKTQCLDANECEAKP 420
QY 101 CLEGECKTLPDN-----SGWM 117
Db 421 CVNAKCKNLIASYCDCLPGWM 443

RESULT 15
US-09-068-740A-6
Sequence 6, Application US/09068740A
Patent No. 6337387
GENERAL INFORMATION:
APPLICANT: SAKANO, SEIJI
APPLICANT: ITOH, AKIRA
TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
FILE REFERENCE: KP-8447
CURRENT APPLICATION NUMBER: US/09/068,740A
CURRENT FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: JP 7-299611
PRIOR FILING DATE: 1995-11-17
PRIOR APPLICATION NUMBER: JP 7-311811
PRIOR FILING DATE: 1995-11-30
PRIOR APPLICATION NUMBER: PCT/JP96/03356
PRIOR FILING DATE: 1996-11-15
NUMBER OF SEQ ID NOS: 48
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 1036
TYPE: PRT
ORGANISM: Homo sapiens
US-09-068-740A-6

Query Match 11.4%; Score 86; DB 4; Length 1036;
Best Local Similarity 23.8%; Pred. No. 0.75;
Matches 34; Conservative 11; Mismatches 34; Indels 64; Gaps 9;
QY 37 HRP-EGGTCE-----VIAAHR-----C-NKNRIEERSQTVKSC 69
Db 272 HQPLNGGTGTCNTGPKYQSCPEGYSGPNCIAEHAACLSDPCHNRGCKETSLSGFECEC 331
QY 70 LPKGVAGTTNR-----PSCVDASIVIG-----KWW-----CEMEP 100
Db 332 SPGWGTGTCSTNIDDCSPNCSHGTCQD--LVNGFKVCVPPQWTKTQCLDANECEAKP 389
QY 101 CLEGECKTLPDN-----SGWM 117
Db 390 CVNAKCKNLIASYCDCLPGWM 412

Search completed: April 8, 2004, 13:18:40
Job time: 23 secs

A;Description: Sequence of the mouse major histocompatibility locus class III region.
A;Reference number: Z16543
A;Accession: T09059
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1964 <ROW>
A;Cross-references: EMBL:AF030001; NID:g2564945; PID:g2564947
C;Genetics:
A;Gene: notch4
A;Map position: 17
A;Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1; 676/1; 724/1; 772/1
A;Superfamily: notch protein; ankyrin repeat homology; EGF homology
C;Keywords: receptor; signal transduction
F;514-545/Domain: EGF homology <EGF>

Query Match 12.7%; Score 96; DB 2; Length 1964;
Best Local Similarity 26.8%; Pred. No. 0.34;
Matches 33; Conservative 8; Mismatches 48; Indels 34; Gaps 6;

QY 15 ISACAMLLC-HGSLQHTFQOHHLRPEGGTC-----EVTAAHR--CCNKNRI 58
Db 689 LGGCIPTFCAGGTCHP-----QPSGYNCTCPAGYMGTLCTSEEVTAHSGPCLNGGSC 741

QY 59 EERSQTVKCSCLPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLDPNSGWC 118
Db 742 SIPEGYSCITCLP---SHTGRHCQTAVD-----HCVSASCLNGTGVNKPETFFCLC 790

QY 119 ATG 121
Db 791 ATG 793

RESULT 3
S18188
notch protein homolog - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 02-Aug-2002
C;Accession: S18188
R;Weinmaster, G.; Roberts, V.J.; Lemke, G.
Development 113, 199-205, 1991
A;Title: A homolog of Drosophila Notch expressed during mammalian development.
A;Reference number: S18188; MUID:92111383; PMID:1764995
A;Accession: S18188
A;Molecule type: mRNA
A;Residues: 1-2531 <WEI>
A;Cross-references: EMBL:X57405; NID:g57634; PID:g57635
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
F;987-1018/Domain: EGF homology <EGF1>
F;1025-1056/Domain: EGF homology <EGF>
F;1233-1264/Domain: EGF homology <EGF2>
F;1917-1949/Domain: ankyrin repeat homology <AN1>
F;1950-1982/Domain: ankyrin repeat homology <AN2>
F;1984-2016/Domain: ankyrin repeat homology <AN3>
F;2017-2049/Domain: ankyrin repeat homology <AN4>
F;2050-2082/Domain: ankyrin repeat homology <AN5>

Query Match 12.7%; Score 95.5; DB 2; Length 2531;
Best Local Similarity 26.8%; Pred. No. 0.49;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

QY 15 ISACAMLLCHGSLQHTFQOHHLRPEGGTCF-VIAAHR-----CNKN 56
Db 679 IDECAGSPCH-----NGTCTEDGAGTCRCPGYHDPCLSEVNECNSN 723

QY 57 -----RIERSQTVKCSCLPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLTP 111
Db 724 PCIHGACRDGLNGYKCDAPG-WSGTNCD-----INNNECESNCPVNGGTCKDM- 771

QY 112 DNSGWC 118
Db 772 -TSGYVC 777

RESULT 4

A46019
notch-1 protein - mouse
N;Alternate names: notch protein
C;Species: Mus musculus (house mouse)
C;Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 07-Mar-2003
C;Accession: A46019; S25144; C49175; B46438; A46438; PH1569; S32109
R;del Amo, F.F.; Gendron-Maguire, M.; Swiatek, P.J.; Jenkins, N.A.; Copeland, N.G.; Grid Genomics 15, 259-264, 1993
A;Title: Cloning, analysis, and chromosomal localization of Notch-1, a mouse homolog of Drosophila Notch.
A;Reference number: A46019; MUID:93194170; PMID:8449489
A;Accession: A46019
A;Status: not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-2531
A;Cross-references: GB:Z11886; GB:S47228; NID:g288502; PIDN:CAA77941.1; PID:g288503
A;Note: sequence extracted from NCBI backbone (NCBIP:127318)
R;Franco del Amo, F.; Smith, D.E.; Swiatek, P.J.; Gendron-Maguire, M.; Greenspan, R.J.; submitted to the EMBL Data Library, April 1992
A;Description: Expression pattern of Notch, a mouse homolog of Drosophila Notch, suggest
A;Reference number: S25144
A;Accession: S25144
A;Molecule type: mRNA
A;Residues: 1551-2108, 'Q', 2110-2114, 'ALP', 2118-2170 <FRA>
A;Cross-references: EMBL:Z11886
R;Lardelli, M.; Lendahl, U.
Exp. Cell Res. 204, 364-372, 1993
A;Title: Notch A and Notch B--two mouse Notch homologues coexpressed in a wide variety o
A;Reference number: A49175; MUID:93178563; PMID:8440332
A;Accession: C49175
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1161-1547 <LAR>
A;Cross-references: EMBL:X68278; NID:g287987; PIDN:CAA48339.1; PID:g287988
A;Experimental source: embryo
A;Note: sequence extracted from NCBI backbone (NCBIP:126159)
R;Kopan, R.; Weintraub, H.
J. Cell Biol. 121, 631-641, 1993
A;Title: Mouse notch: expression in hair follicles correlates with cell fate determinati
A;Reference number: A46438; MUID:93252998; PMID:8486742
A;Accession: B46438
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1865-1932, 'RR', 1935-1937, 'L', 1938-1967, 'I', 1969-2044, 'IE', 2047-2052, 'S', 2054
A;Experimental source: embryo
A;Note: sequence extracted from NCBI backbone (NCBIN:131246, NCBIP:131247)
C;Comment: This protein has many EGF repeats and lin-12[1172]/Notch repeats.
C;Genetics:
A;Gene: notch-1
A;Map position: 2
A;Note: proximal region of chromosome 2
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
F;106-138/Domain: EGF homology <EGF1>
F;144-175/Domain: EGF homology <EGO1>
F;222-254/Domain: EGF homology <EGF2>
F;261-292/Domain: EGF homology <EGO2>
F;339-370/Domain: EGF homology <EGO3>
F;416-449/Domain: EGF homology <EGF3>
F;456-487/Domain: EGF homology <EGO4>
F;494-525/Domain: EGF homology <EGO5>
F;532-563/Domain: EGF homology <EGO6>
F;607-638/Domain: EGF homology <EGO7>
F;682-713/Domain: EGF homology <EGO8>
F;757-788/Domain: EGF homology <EGO9>
F;795-826/Domain: EGF homology <EGO10>
F;873-904/Domain: EGF homology <EGO11>
F;911-942/Domain: EGF homology <EGO12>
F;949-980/Domain: EGF homology <EGO13>
F;987-1018/Domain: EGF homology <EGO14>
F;1025-1056/Domain: EGF homology <EGO15>
F;1063-1094/Domain: EGF homology <EGO16>

F;1149-1180/Domain: EGF homology <EG17>
F;1187-1218/Domain: EGF homology <EG18>
F;1233-1264/Domain: EGF homology <EGF4>
F;1352-1383/Domain: EGF homology <EGF9>
F;1391-1425/Domain: EGF homology <EGF>
F;1917-1948/Domain: ankyrin repeat homology <AN1>
F;1949-1981/Domain: ankyrin repeat homology <AN2>
F;1983-2015/Domain: ankyrin repeat homology <AN3>
F;2016-2048/Domain: ankyrin repeat homology <AN4>
F;2049-2081/Domain: ankyrin repeat homology <AN5>

Query Match 12.7%; Score 95.5; DB 2; Length 2531;
Best Local Similarity 26.8%; Pred. No. 0.49;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

QY 15 ISACAMLLCHGSLQHTFQHHLRPEGTCE-VIAAHR-----CNKN 56
Db 679 IDECAGSPCH-----NGTCDGIAGFTCRCPGYPHPTCLSEVNECNSN 723
QY 57 -----RIEERSQTVKCSCLPGKVGAGTTRNRPSQVDASIVIGKWCMEPCLEGECKTLP 111
Db 724 PCIHGACRDGLNGYKCDAPG-WSGTNCD-----INNNECESNPNVNGGTCKDM- 771
QY 112 DNSGWMC 118
Db 772 -TSGYVC 777

RESULT 5

T32021 Notch homolog protein - sea squirt (Halocynthia roretzi)

C;Species: Halocynthia roretzi
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 08-Sep-2002
C;Accession: T32021
R;Hori, S.; Saiboh, T.; Matsumoto, M.; Makabe, K.W.; Nishida, H.
Dev. Genes Evol. 207, 371-380, 1997
A;Title: Notch homologue from Halocynthia roretzi is preferentially expressed in the central nervous system
A;Reference number: Z20775
A;Accession: T32021
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2352 <HOR>
A;Cross-references: EMBL:AB001327; NID:dl204472; PID:dl026501; PIDN:BA25571.1
C;Genetics:
A;Gene: Notch
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

Query Match 12.5%; Score 94.5; DB 2; Length 2352;
Best Local Similarity 26.0%; Pred. No. 0.57;
Matches 27; Conservative 12; Mismatches 34; Indels 31; Gaps 5;

QY 32 QQHHLRPEGG-TCEVIAAH-----RCNKNRIEERSQTVKCSCLPGKVA 75
Db 236 KRRDLQTEGGFTCNVYGFTRDDCSENIIDCSNVACFHNARCIDQAGTFECLCTPG--- 292
QY 76 GTTNRPSQ-VDAIVIGKWCMEPCLEGECKTLPNSGWMC 118
Db 293 ---NRILCHLDDA-----CISDPGARGATCDTNPITGHWC 325

RESULT 6

T31070

Notch homolog - sea urchin (Lytechinus variegatus)
C;Species: Lytechinus variegatus (variegated urchin)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C;Accession: T31070
R;Sherwood, D.R.; McClay, D.R.
Development 124, 3363-3374, 1997
A;Title: Identification and localization of a sea urchin Notch homologue: insights into the evolution of the Notch signaling pathway
A;Reference number: Z20966; MUID:97454256; PMID:9310331
A;Accession: T31070
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA

A;Residues: 1-2531 <SHE>
A;Cross-references: EMBL:AF000634; NID:g2570350; PID:g2570351; PIDN:AAB82088.1
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

Query Match 12.3%; Score 93; DB 2; Length 2531;
Best Local Similarity 22.4%; Pred. No. 0.85;
Matches 28; Conservative 18; Mismatches 41; Indels 38; Gaps 5;

QY 40 EGGTCEV-----IAAHRCCNKNRIEERSQTVKCSCLPGKVGAGTTRNRPSQV----- 85
Db 350 EGGTCEINKDDCTPNPCQPEGECEDRVASPKCTCPGRTGLLCHLEDACMSNPFCHHTAQ 409
QY 86 DASIVIGKWCMEPCLEGE-----ECKTLPD-----NSGWMCA-----TG 121
Db 410 STSVVDGSGFTCDATGYQGFNCSEDIDECSSLMSDSICQSGGTCTQNFQDGSCLCSGFTG 469
QY 122 NKIKT 126
Db 470 SRET 474

RESULT 7

A35672

crumbs protein - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster
C;Date: 21-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 21-Jul-2003
C;Accession: A35672
R;Tepass, U.; Theres, C.; Knust, E.
Cell 61, 787-799, 1990
A;Title: crumbs encodes an EGF-like protein expressed on apical membranes of Drosophila
A;Reference number: A35672; MUID:90263104; PMID:2344615
A;Accession: A35672
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2139 <TEP>
A;Cross-references: GB:M33753
A;Note: the authors translated the codon GGC for residue 1928 as Cys, and TAT for residue 1929 as Phe
C;Genetics:
A;Gene: FlyBase:crb
A;Cross-references: FlyBase:FBgn0000368
C;Keywords: transmembrane protein
F;352-385/Domain: EGF homology <EGX1>
F;392-424/Domain: EGF homology <EGF1>
F;691-722/Domain: EGF homology <EGF>
F;767-799/Domain: EGF homology <EGF3>
F;1878-1914/Domain: EGF homology <EGX2>

Query Match 12.3%; Score 92.5; DB 2; Length 2139;
Best Local Similarity 26.0%; Pred. No. 0.82;
Matches 27; Conservative 14; Mismatches 46; Indels 17; Gaps 5;

QY 41 GGTCEVIAAHRCC-----NKNRIEERSQTVKCSCLPGKVGAGTTRNRPSQVDASIVIGKWC 96
Db 559 GGTCSMNGTHCYCAVGYSGDRCEKAENCSPLNCPQEPWVC--VQNCCLCPENKVCN---QC 613
QY 97 EMEPCLEGECKTLPDNS-----GWCAT-GNKIKTTIHRP 132
Db 614 ATPCQNGGECVDLPNGDYECKTRGTGTCGNDVDDECTLHPK 657

RESULT 8

A24420

Notch protein - fruit fly (Drosophila melanogaster)
N;Alternate names: neurogenic repetitive locus protein
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A24420; A24768; S09358; A05267
R;Kidd, S.; Kelley, M.R.; Young, M.W.
Mol. Cell. Biol. 6, 3094-3108, 1986
A;Reference number: A24420; MUID:87064624; PMID:3097517
A;Accession: A24420
A;Molecule type: DNA
A;Residues: 1-2703 <KID>

A;Cross-references: GB:K03508; NID:G157991; PIDN:AAA28725.1; PID:G157993
 R;Wharton, K.A.; Johansen, K.M.; Xu, T.; Artavanis-Tsakonas, S.
 Cell 43, 567-581, 1985
 A;Reference number: A24768; MUID:86079539; PMID:3935325
 A;Accession: A24768
 A;Molecule type: mRNA
 A;Residues: 1-48, 'I', 50-118, 'R', 120-230, 'I', 232-256, 'N', 258-266, 'A', 268-872, 'R', 874-958,
 A;Note: the authors translated the codon ATC for residue 49 as Thr, ATT for residue 2044
 R;Tautz, D.
 Nucleic Acids Res. 17, 6463-6471, 1989
 A;Title: Hypervariability of simple sequences as a general source for polymorphic DNA ma
 A;Reference number: S09358; MUID:89385974; PMID:2780284
 A;Accession: S09358
 A;Molecule type: DNA
 A;Residues: 2505-2551, 'QQQ', 2552-2576, 'E', 2578-2604 'TAU'
 R;Wharton, K.A.; Yedvobnick, B.; Finnerty, V.G.; Artavanis-Tsakonas, S.
 Cell 40, 55-62, 1985
 A;Title: Opa: a novel family of transcribed repeats shared by the Notch locus and other
 A;Reference number: A05267; MUID:85099329; PMID:2981631
 A;Accession: A05267
 A;Molecule type: DNA
 A;Residues: 2504-2576, 'E', 2578-2611 'WHA2'
 C;Genetics:
 A;Gene: notch; opa
 A;Cross-references: FlyBase:FBgn0004647
 A;Map position: 8.96-9.36
 A;Introns: 53/3; 84/3; 171/3; 240/3; 283/3; 2333/3; 2436/3; 2588/3
 C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
 C;Keywords: differentiation; tandem repeat; transmembrane protein
 F;27-43/Domain: transmembrane #status predicted <TMW1>
 F;297-328/Domain: EGF homology <EGX1>
 F;530-561/Domain: EGF homology <EGF1>
 F;568-599/Domain: EGF homology <EGF>
 F;988-1019/Domain: EGF homology <EGX2>
 F;1064-1095/Domain: EGF homology <EGF3>
 F;1187-1218/Domain: EGF homology <EGX3>
 F;1746-1762/Domain: transmembrane #status predicted <TMW2>
 F;1950-1982/Domain: ankyrin repeat homology <AN1>
 F;1983-2015/Domain: ankyrin repeat homology <AN2>
 F;1988-2004/Domain: transmembrane #status predicted <TMW3>
 F;2017-2049/Domain: ankyrin repeat homology <AN3>
 F;2050-2082/Domain: ankyrin repeat homology <AN4>
 F;2083-2115/Domain: ankyrin repeat homology <AN5>
 F;2538-2568/Region: glutamine-rich
 F;2538-2568/Domain: neurogenic repetitive element #status predicted <OPA>
 Query Match 11.5%; Score 86.5; DB 1; Length 2703;
 Best Local Similarity 25.9%; Pred. No. 3.9;
 Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;
 QY 40 EGGTCEV---IAHRCCNKNRIEERSQTVKCSCLPGKVA-----GTTNRPS 83
 Db 1214 QGONCELINDDCAPNQCNGTGRVNFSCPPGTGIIKNDCKPGACHNGS 1273
 QY 84 CVDASIVIGKWCMEPCLEGECK 108
 Db 1274 CIDR---VGGFECVQPGFVGARCE 1295
 RESULT 9
 JC7570
 Delta-4 protein - human
 C;Species: Homo sapiens (man)
 C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 16-Aug-2002
 R;Yoneya, T.; Tahara, T.; Nagao, K.; Yamada, Y.; Yamamoto, T.; Osawa, M.; Miyatani, S.;
 J. Biochem. 129, 27-34, 2001
 A;Title: Molecular cloning of Delta-4, a new mouse and human notch ligand.
 A;Reference number: JC7569; MUID: 21064937; PMID:11134954
 A;Accession: JC7570
 A;Molecule type: mRNA
 A;Residues: 1-685 <YON>
 A;Cross-references: DDBJ:AB043894

C;Comment: This protein, a member of the Notch family of proteins, is a transmembrane re
 ates the Notch signaling, the growth or differentiation of vascular endothelial cells.
 C;Genetics:
 A;Gene: delta-4
 C;Superfamily: delta-4 protein; EGF homology
 C;Keywords: transmembrane protein
 Query Match 11.4%; Score 86; DB 2; Length 685;
 Best Local Similarity 25.0%; Pred. No. 1.2;
 Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;
 QY 14 WISAC-----AMLLCHSLOHTFOOHLHPEGTCEVIAHRCCKNRIEERSQTVKCS 69
 Db 268 WQCTDEGWLFCDDQLNYC-----THSPCKNGATC-----SNSGQSYT--CTC 312
 QY 70 LPQKVAAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDSCMWCATG 121
 Db 313 RPYGTG-----VDCELEISE--CDNPNCRNGSGCKDQEDGYHCLPPG 353
 RESULT 10
 A56136
 jagged protein precursor - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 21-Jul-2003
 C;Accession: A56136
 R;Kindell, C.E.; Shawber, C.J.; Boulter, J.; Weinmaster, G.
 Cell 80, 909-917, 1995
 A;Title: Jagged: a mammalian ligand that activates Notchl.
 A;Reference number: A56136; MUID:95211842; PMID:7697721
 A;Accession: A56136
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-1220 <LIN>
 A;Cross-references: GB:I138483
 F;379-410/Domain: EGF homology <EGF1>
 F;492-523/Domain: EGF homology <EGF>
 F;634-665/Domain: EGF homology <EGF2>
 Query Match 11.4%; Score 86; DB 2; Length 1220;
 Best Local Similarity 27.0%; Pred. No. 2.1;
 Matches 33; Conservative 9; Mismatches 34; Indels 46; Gaps 9;
 QY 39 PE---GCTCEVIAHRC---C-NKNRIEERSQTVKCSCLPGKVAAGTTRNR----- 81
 Db 326 PEGYSGPNCE-IAEHACLSDPCHNRGSKETSGPECECSGPGWTGPTCTNIDDCSPNNC 384
 QY 82 ---PSCVDASIVIG-----KWW-----CEMEPCLEGECKTLPD-----SG 115
 Db 385 SHGGTCQD--LVNGFKVCVPPQWTGKTCQLDANECAKFCVNARSKNLIASYCDCLPG 442
 QY 116 WM 117
 Db 443 WM 444
 RESULT 11
 A40136
 fibropellin Ia - sea urchin (Strongylocentrotus purpuratus)
 N;Alternate names: epidermal growth factor homolog precursor
 N;Contains: alternatively spliced fibropellin Ib (EGPI)
 C;Species: Strongylocentrotus purpuratus (purple urchin)
 C;Date: 13-May-1992 #sequence_revision 17-Sep-1997 #text_change 21-Jul-2000
 C;Accession: A40136; B40136; C40136; A29316; A43131
 R;Delgadillo-Reynoso, M.G.; Rollo, D.R.; Hursh, D.A.; Raff, R.A.
 J. Mol. Evol. 29, 314-327, 1989
 A;Title: Structural analysis of the uEGF gene in the sea urchin Strongylocentrotus purpu
 A;Reference number: A40136; MUID:90112459; PMID:2514273
 A;Accession: A40136
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-114
 A;Cross-references: GB:X17530; NID:G10225; PID:G667061

```
Query Match      10.9%; Score 82.5; DB 2; Length 1203;
Best Local Similarity 26.4%; Pred. No. 4.5;
Matches 23; Conservative 11; Mismatches 36; Indels 17; Gaps 4;

Qy      40 EGGTCFV-----IAAHRCCNKNRIEERSQTWKCSCLPGKVAGTTNRPS-CVDASIVIGKW 94
```

Db 170 KGVHCELVNECQSNPCVNNNGQCVKVNRFQCLCPGFTG-----PVCQIDIDD----- 218

Qy 95 WCMEPCLEGECKTLDPDNGWMCATG 121

Db 219 -CSSTPCLINGAKCIDHPNGYEQCATG 244

RESULT 14

A49128 cell-fate determining gene Notch2 protein - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 02-Aug-2002

C;Accession: A49128

R;Weinmaster, G.; Roberts, V.J.; Lemke, G.

Development 116, 931-941, 1992

A;Title: Notch2: a second mammalian Notch gene.

A;Reference number: A49128; MUID:93202015; PMID:1295745

A;Accession: A49128

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-2471 <WEI>

A;Experimental source: Schwann cell

A;Note: sequence extracted from NCBI backbone (NCBIP:127811)

C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

F;264-295/Domain: EGF homology <EGX1>

F;799-830/Domain: EGF homology <EGF1>

F;877-908/Domain: EGF homology <EGX2>

F;1029-1060/Domain: EGF homology <EGF>

F;1067-1098/Domain: EGF homology <EGX3>

F;1153-1184/Domain: EGF homology <EGF3>

F;1191-1222/Domain: EGF homology <EGX4>

F;1876-1908/Domain: ankyrin repeat homology <AN1>

F;1909-1941/Domain: ankyrin repeat homology <AN2>

F;1943-1975/Domain: ankyrin repeat homology <AN3>

F;1976-2008/Domain: ankyrin repeat homology <AN4>

F;2009-2041/Domain: ankyrin repeat homology <AN5>

Query Match 10.9%; Score 82.5; DB 2; Length 2471;

Best Local Similarity 26.4%; Pred. No. 8.7;

Matches 23; Conservative 11; Mismatches 36; Indels 17; Gaps 4;

Qy 40 EGGTCEV-----IAHRCCKNRIERSQTVKCSCLPGKAVGTTNRPS-CVDASIVIGKW 94

Db 487 KGVHCELVNECQSNPCVNNNGQCVKVNRFQCLCPGFTG-----PVCQIDIDD----- 535

Qy 95 WCMEPCLEGECKTLDPDNGWMCATG 121

Db 536 -CSSTPCLINGAKCIDHPNGYEQCATG 561

RESULT 15

S45306

notch 3 protein - mouse

C;Species: Mus musculus (house mouse)

C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Aug-2002

C;Accession: S45306

R;Lardelli, M.; Dahlstrand, J.; Lendahl, U.

Mech. Dev. 46, 123-136, 1994

A;Title: The novel Notch homologue mouse Notch 3 lacks specific epidermal growth factor-

A;Reference number: S45306; MUID:95001556; PMID:7918097

A;Accession: S45306

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-2318 <LAR>

A;Cross-references: EMBL:X74760; NID:9483580; PIDN:CAA52776.1; PID:9483581

C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

F;163-195/Domain: EGF homology <EGF1>

F;474-505/Domain: EGF homology <EGF>

F;854-885/Domain: EGF homology <EGF2>

F;1839-1871/Domain: ankyrin repeat homology <AN1>

F;1872-1904/Domain: ankyrin repeat homology <AN2>

F;1906-1938/Domain: ankyrin repeat homology <AN3>

F;1939-1971/Domain: ankyrin repeat homology <AN4>

F;1972-2004/Domain: ankyrin repeat homology <AN5>

Query Match 10.9%; Score 82; DB 2; Length 2318;

Best Local Similarity 22.8%; Pred. No. 9.2;

Matches 39; Conservative 16; Mismatches 42; Indels 74; Gaps 10;

Qy 17 ACAM-----LLCHGSLQHTFQQHHLHRP-----EGGTCE-----V 46

Db 340 ACPMGKTGLLCH--LDDACVSNPCHEDAICDTNPFVSGRAICTCPGFTGACDQDVDECS 397

Qy 47 IAAHRCCKNRIERSQTVKCSCLPGK-----VAGTTRNRPSCVDASIVIGKW 94

Db 398 IGANPCEHLGRCVNTGCSFLCCQGRGYTGPRCETDVNECLSGPCRQATCLDR---IGQF 454

Qy 95 -----WCME-----PCLEGECK-----TLFDN-SGMMC 118

Db 455 TCICMAGFTGTYCEVIDIDECQSSPCVNGGVCKDRVNGFSTCFSGFSGSMC 505

Search completed: April 8, 2004, 13:18:05

Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 13:17:35 ; Search time 40 Seconds
(without alignments)
874.269 Million cell updates/sec

Title: US-09-763-335-2

Perfect score: 754
Sequence: 1 MAMVSAMSVMVLYLWISACAM.....SGWMCATGNKIKTRHPRT 133

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1073127 seqs, 262937947 residues

Total number of hits satisfying chosen parameters: 1073127

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pap.*
- 2: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pap.*
- 3: /cgn2_6/ptodata/1/pubpa/US06_NEW_PUB.pap.*
- 4: /cgn2_6/ptodata/1/pubpa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/1/pubpa/US07_NEW_PUB.pap.*
- 6: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/1/pubpa/US08_NEW_PUB.pap.*
- 8: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pap.*
- 10: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pap.*
- 12: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pap.*
- 13: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pap.*
- 14: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pap.*
- 15: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pap.*
- 16: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pap.*
- 17: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pap.*
- 18: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	14.1	3871	15	US-10-369-493-6538
2	106	14.1	3871	15	US-10-369-493-6538
3	95.5	12.7	2531	12	US-10-072-012-470
4	95.5	12.7	2531	12	US-10-072-012-471
5	95.5	12.7	2531	15	US-10-190-115-29
6	95.5	12.7	2531	15	US-10-369-072-29
7	91.5	12.1	233	9	US-09-216-393-110
8	91.5	12.1	233	14	US-10-321-856-110
9	88.5	11.7	2447	15	US-10-190-115-28
10	88.5	11.7	2447	15	US-10-369-072-28
11	87	11.5	188	9	US-09-764-853-773
12	87	11.5	188	9	US-09-764-898-262
13	87	11.5	188	10	US-09-764-881-109
14	87	11.5	188	14	US-10-073-865-88
15	87	11.5	188	15	US-10-242-747-109

16	86.5	11.5	1139	12	US-10-419-026-4	Sequence 4, Appli
17	86.5	11.5	2556	12	US-10-072-012-134	Sequence 134, App
18	86	11.4	174	9	US-09-764-898-189	Sequence 189, App
19	86	11.4	500	14	US-10-241-476-2	Sequence 2, Appli
20	86	11.4	528	14	US-10-225-630-4	Sequence 4, Appli
21	86	11.4	528	14	US-10-241-476-3	Sequence 3, Appli
22	86	11.4	559	14	US-10-147-493-88	Sequence 88, Appl
23	86	11.4	585	12	US-10-145-127-88	Sequence 88, Appl
24	86	11.4	685	12	US-10-160-503-88	Sequence 88, Appl
25	86	11.4	685	12	US-10-143-118-88	Sequence 88, Appl
26	86	11.4	685	12	US-10-144-993-88	Sequence 88, Appl
27	86	11.4	685	12	US-10-158-787-88	Sequence 88, Appl
28	86	11.4	685	12	US-10-081-056-276	Sequence 276, App
29	86	11.4	685	12	US-10-219-535-214	Sequence 214, App
30	86	11.4	685	12	US-10-232-230-214	Sequence 214, App
31	86	11.4	685	12	US-10-140-024-88	Sequence 88, Appl
32	86	11.4	685	14	US-10-028-072-88	Sequence 88, Appl
33	86	11.4	685	14	US-10-121-049-88	Sequence 88, Appl
34	86	11.4	685	14	US-10-123-904-88	Sequence 88, Appl
35	86	11.4	685	14	US-10-140-470-88	Sequence 88, Appl
36	86	11.4	685	14	US-10-175-746-88	Sequence 88, Appl
37	86	11.4	685	14	US-10-176-918-88	Sequence 88, Appl
38	86	11.4	685	14	US-10-176-921-88	Sequence 88, Appl
39	86	11.4	685	14	US-10-227-884-214	Sequence 214, App
40	86	11.4	685	14	US-10-137-865-88	Sequence 88, Appl
41	86	11.4	685	14	US-10-140-474-88	Sequence 88, Appl
42	86	11.4	685	14	US-10-142-431-88	Sequence 88, Appl
43	86	11.4	685	14	US-10-143-114-88	Sequence 88, Appl
44	86	11.4	685	14	US-10-230-163-214	Sequence 214, App
45	86	11.4	685	14	US-10-140-002-88	Sequence 88, Appl

ALIGNMENTS

RESULT 1

US-10-369-493-6538
; Sequence 6538, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6538
; LENGTH: 3871
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6538

Query Match 14.1%; Score 106; DB 15; Length 3871;
Best Local Similarity 31.2%; Pred. No. 0.15;
Matches 34; Conservative 13; Mismatches 40; Indels 22; Gaps 6;
QY 24 HGSQHTFQOHUHLRREGGTCEVIAAHRCKNKNRIERSQ-----TVKCSLPGKVGAT 77
Db 148 HGSLL-----CQYHMSAC-SKTFELCGPHGHCIESIVDPTGSSSDTTTTCIKICDWGFKVSS 202
QY 78 TRNRSPCVDAISIVIGKWCMECECLGECKTLPDN---SGWMCATGNK 123
Db 203 DKNNTCTVDVN-----ECESNFCHPGVDCINLPFGSVCSG--CPKGYK 243
RESULT 2

```

; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 470
; . LENGTH: 2531
; . TYPE: PRT
; . ORGANISM: Rattus norvegicus
;
US-10-072-012-470

Query Match          12.7%; Score 95.5; DB 12; Length 2531;
Best Local Similarity 26.8%; Pred. No. 1.2;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

QY      15  ISACMLICHGSLQHTFQOHLHLRPEGGTCE-VIAAHC-----CNKN 56
Db      679  IDECAQSPCH-----NGGTCEDGIAGFTCRCPGYHDPTCLSEVNECNSN 723

QY      57  -----RIERSQTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
Db      724  PCIHGACRDGLNGYKDCAPG-WSGTNCD-----INNNECESNPCVNGTCKDM- 471

QY      112  DNSGMMC 118
Db      772  -TSGYVC 777

RESULT 4
US-10-072-012-471
; Sequence 471, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Ganggoli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Teupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10072,012
; CURRENT FILING DATE: 2002-01-31

```

```

; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 471
; LENGTH: 2531
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-072-012-471

```

```

Query Match 12.7%; Score 95.5; DB 12; Length 2531;
Best Local Similarity 26.8%; Pred. No. 1.2;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

QY 15 ISACMLLCHGSLQHTFQOHLHRPEGTCE-VIAAHR-----CNKN 56
Db 679 IDECAGSPCH-----NGGTCDGAGTCTCRCEGYHDPCTCLSEVNECNSN 723

QY 57 -----RIERSQTVKCSCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
Db 724 PCHGACRDLNGYKDCAPG-WSGTNCD-----INNNECESNPCVNGGTCKDM- 771

QY 112 DNSGMMC 118
Db 772 -TSGYVC 777

```

```

RESULT 5
US-10-190-115-29
; Sequence 29, Application US/10190115
; Publication No. US20030207394A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, John P. II
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Grosse, William M.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Lepley, Denise M.
; APPLICANT: Liu, Xiaohong
; APPLICANT: Mezzick, Amanda J.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Rastelli, Luca
; APPLICANT: Shen, Lei
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Szekeres, Edward S. Jr.
; APPLICANT: Taupier, Raymond J. Jr.
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Zerhusen, Bryan D.

```

```

; APPLICANT: Voss, Edward Z.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-050 CIP
; CURRENT APPLICATION NUMBER: US/10/190,115
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/303,168
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585,
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2001-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 29
; LENGTH: 2531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-190-115-29

```

```

Query Match 12.7%; Score 95.5; DB 15; Length 2531;
Best Local Similarity 26.8%; Pred. No. 1.2;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

QY 15 ISACMLLCHGSLQHTFQOHLHRPEGTCE-VIAAHR-----CNKN 56
Db 679 IDECAGSPCH-----NGGTCDGAGTCTCRCEGYHDPCTCLSEVNECNSN 723

QY 57 -----RIERSQTVKCSCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
Db 724 PCHGACRDLNGYKDCAPG-WSGTNCD-----INNNECESNPCVNGGTCKDM- 771

QY 112 DNSGMMC 118
Db 772 -TSGYVC 777

```

```

RESULT 6
US-10-369-072-29
; Sequence 29, Application US/10369072
; Publication No. US20040014081A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szekeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara

```

; TITLE OF INVENTION: No. US20040014081a1el Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-050 CON2
 ; CURRENT APPLICATION NUMBER: US/10/369,072
 ; CURRENT FILING DATE: 2003-02-18
 ; PRIOR APPLICATION NUMBER: 10/174,372
 ; PRIOR FILING DATE: 2002-06-17
 ; PRIOR APPLICATION NUMBER: 09/898,994
 ; PRIOR FILING DATE: 2001-07-03
 ; PRIOR APPLICATION NUMBER: 60/215,854
 ; PRIOR FILING DATE: 2000-07-03
 ; PRIOR APPLICATION NUMBER: 60/215,856
 ; PRIOR FILING DATE: 2000-07-03
 ; PRIOR APPLICATION NUMBER: 60/215,902
 ; PRIOR FILING DATE: 2000-07-03
 ; PRIOR APPLICATION NUMBER: 60/216,585
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/216,586
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/216,722
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/218,622
 ; PRIOR FILING DATE: 2000-07-17
 ; PRIOR APPLICATION NUMBER: 60/218,992
 ; PRIOR FILING DATE: 2000-07-17
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 100
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 29
 ; LENGTH: 2531
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 ; ORGANISM: Rattus norvegicus
 US-10-369-072-29

Query Match 12.7%; Score 95.5; DB 15; Length 2531;
 Best Local Similarity 26.8%; Pred. No. 1.2;
 Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;
 QY 15 ISACMLLCHGSLQHTFQOHLHLPREGTCE-VIAAHR-----CNKN 56
 DB 679 IDECAGSPCH-----NGTCEGDIAGTCTCPGEGYDPTCLSEVNECNSN 723
 QY 57 -----RIEERSQTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
 DB 724 PCIHGACRDGLNGYKCDAPG-WSGTNCD-----INNNECESNPVNGGTCKDM- 771
 QY 112 DNSGWC 118
 DB 772 -TSGYVC 777

RESULT 7
 US-09-216-393-110
 ; Sequence 110, Application US/09216393
 ; Patent No. US20010014447A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Milhausen, Michael James
 ; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
 ; FILE REFERENCE: TX-1-C2
 ; CURRENT APPLICATION NUMBER: US/09/216,393
 ; CURRENT FILING DATE: 1998-12-18
 ; EARLIER APPLICATION NUMBER: 08/994,825
 ; EARLIER FILING DATE: 1997-12-19
 ; NUMBER OF SEQ ID NOS: 364
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 110
 ; LENGTH: 233
 ; TYPE: PRT
 ; ORGANISM: Toxoplasma gondii
 US-09-216-393-110

Query Match 12.1%; Score 91.5; DB 9; Length 233;
 Best Local Similarity 25.9%; Pred. No. 0.23;

Matches 28; Conservative 6; Mismatches 39; Indels 35; Gaps 4;
 QY 18 CAMLLCHGSLQHTFQOHLHLPREGTCEVIAAHR-----CNKN-----RIEERS 62
 DB 78 QCQDACHG-----GKTCEITKEHCCINDSDCNHGHGTCTNSNTNCEAGF 122
 QY 63 QTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLP 110
 DB 123 AGTNCSSSEKCSGKT-----CLSGHCNPATGACVCDPCHTGERCETL 165
 RESULT 8
 US-10-321-856-110
 ; Sequence 110, Application US/10321856
 ; Publication No. US20030194393A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Milhausen, Michael James
 ; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
 ; FILE REFERENCE: TX-1-C2-1
 ; CURRENT APPLICATION NUMBER: US/10/321,856
 ; CURRENT FILING DATE: 2002-12-17
 ; PRIOR APPLICATION NUMBER: 09/216,393
 ; PRIOR FILING DATE: 1998-12-18
 ; PRIOR APPLICATION NUMBER: 08/994,825
 ; PRIOR FILING DATE: 1997-12-19
 ; NUMBER OF SEQ ID NOS: 366
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 110
 ; LENGTH: 233
 ; TYPE: PRT
 ; ORGANISM: Toxoplasma gondii
 US-10-321-856-110

Query Match 12.1%; Score 91.5; DB 14; Length 233;
 Best Local Similarity 25.9%; Pred. No. 0.23;
 Matches 28; Conservative 6; Mismatches 39; Indels 35; Gaps 4;
 QY 18 CAMLLCHGSLQHTFQOHLHLPREGTCEVIAAHR-----CNKN-----RIEERS 62
 DB 78 QCQDACHG-----GKTCEITKEHCCINDSDCNHGHGTCTNSNTNCEAGF 122
 QY 63 QTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLP 110
 DB 123 AGTNCSSSEKCSGKT-----CLSGHCNPATGACVCDPCHTGERCETL 165

RESULT 9
 US-10-190-115-28
 ; Sequence 28, Application US/10190115
 ; Publication No. US20030207394A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alsbrook, John P. II
 ; APPLICANT: Boldog, Ferenc L.
 ; APPLICANT: Burgess, Catherine E.
 ; APPLICANT: Casman, Stacie J.
 ; APPLICANT: Grosse, William M.
 ; APPLICANT: Gusev, Vladimir Y.
 ; APPLICANT: Ji, Weizhen
 ; APPLICANT: Lepley, Denise M.
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Mezick, Amanda J.
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Shen, Lei
 ; APPLICANT: Shenoy, Suresh G.
 ; APPLICANT: Shinkets, Richard A.
 ; APPLICANT: Spaderna, Steven K.
 ; APPLICANT: Sytek, Kimberly A.
 ; APPLICANT: Szekeres, Edward S. Jr.
 ; APPLICANT: Taupier, Raymond J. Jr.
 ; APPLICANT: Tchernev, Velizar T.
 ; APPLICANT: Zerhusen, Bryan D.

```

; APPLICANT: Voss, Edward Z.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-050 CIP
; CURRENT APPLICATION NUMBER: US/10/190,115
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/303,168
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: Curasequid version 0.1
; SEQ ID NO 28
; LENGTH: 2447
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-190-115-28

```

```

Query Match      11.7%; Score 88.5; DB 15; Length 2447;
Best Local Similarity 25.4%; Pred. No. 5.9;
Matches 30; Conservative 10; Mismatches 39; Indels 39; Gaps 5;

QY      41  GGTC-----VIAHRCNKNRIERSQTVKCSCLPGKVAGT-----TRNRPS 83
Db      404  GPRCEQDVNECASNPQNDGTCLDRIGDYSCICMPG-FGGTHCENELNECLSSPCLNRGK 462

QY      84  CVDSASIVIGKWCE-----MPCLEGECKTLPDMSGMCMCATGNK 123
Db      463  CLDQ----VSRFVCECPAGFSGEMCQIDIDECSTPCLNGAKCIDLPNGYDCECAEGFK 517

```

```

RESULT 10
US-10-369-072-28
; Sequence 28, Application US/10369072
; Publication No. US20040014081A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkes, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: No. US20040014081A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18

```

```

; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 2447
; TYPE: PRT
; ORGANISM: Takifugu rubripes
US-10-369-072-28

```

```

Query Match      11.7%; Score 88.5; DB 15; Length 2447;
Best Local Similarity 25.4%; Pred. No. 5.9;
Matches 30; Conservative 10; Mismatches 39; Indels 39; Gaps 5;

QY      41  GGTC-----VIAHRCNKNRIERSQTVKCSCLPGKVAGT-----TRNRPS 83
Db      404  GPRCEQDVNECASNPQNDGTCLDRIGDYSCICMPG-FGGTHCENELNECLSSPCLNRGK 462

QY      84  CVDSASIVIGKWCE-----MPCLEGECKTLPDMSGMCMCATGNK 123
Db      463  CLDQ----VSRFVCECPAGFSGEMCQIDIDECSTPCLNGAKCIDLPNGYDCECAEGFK 517

```

```

RESULT 11
US-09-764-853-773
; Sequence 773, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 773
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (187)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (188)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-773

Query Match      11.5%; Score 87; DB 9; Length 188;
Best Local Similarity 25.0%; Pred. No. 0.53;

```

```
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;
QY 14 WISAC-----AMLLCHGSLQHTFOQHLHRRPBGTCVIAAHRCCNKNRIERSQTVKSC 69
Db 18 WQCTDEGWGGLFCDDQLNYC--THSPCKNGATC-----SNSGQRSYT--CTC 62
QY 70 LPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATG 121
Db 63 RPYGTG-----VDCELELSE--CDSNFCRNGGCKDQEDGYXCLCPFG 103

RESULT 12
US-09-764-898-262
; Sequence 262, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ201
; CURRENT APPLICATION NUMBER: US/09/764,898
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 262
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (187)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (188)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-898-262

Query Match 11.5%; Score 87; DB 9; Length 188;
Best Local Similarity 25.0%; Pred. No. 0.53;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;
QY 14 WISAC-----AMLLCHGSLQHTFOQHLHRRPBGTCVIAAHRCCNKNRIERSQTVKSC 69
Db 18 WQCTDEGWGGLFCDDQLNYC--THSPCKNGATC-----SNSGQRSYT--CTC 62
QY 70 LPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATG 121
Db 63 RPYGTG-----VDCELELSE--CDSNFCRNGGCKDQEDGYXCLCPFG 103

RESULT 13
US-09-764-881-109
; Sequence 109, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 109
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (97)
```

```
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (187)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (188)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-881-109

Query Match 11.5%; Score 87; DB 10; Length 188;
Best Local Similarity 25.0%; Pred. No. 0.53;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;
QY 14 WISAC-----AMLLCHGSLQHTFOQHLHRRPBGTCVIAAHRCCNKNRIERSQTVKSC 69
Db 18 WQCTDEGWGGLFCDDQLNYC--THSPCKNGATC-----SNSGQRSYT--CTC 62
QY 70 LPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATG 121
Db 63 RPYGTG-----VDCELELSE--CDSNFCRNGGCKDQEDGYXCLCPFG 103

RESULT 14
US-10-073-865-88
; Sequence 88, Application US/10073865
; Publication No. US2003004904A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ209C1
; CURRENT APPLICATION NUMBER: US/10/073,865
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc feature
; LOCATION: (187)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc feature
; LOCATION: (188)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-073-865-88

Query Match 11.5%; Score 87; DB 14; Length 188;
Best Local Similarity 25.0%; Pred. No. 0.53;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;
QY 14 WISAC-----AMLLCHGSLQHTFOQHLHRRPBGTCVIAAHRCCNKNRIERSQTVKSC 69
Db 18 WQCTDEGWGGLFCDDQLNYC--THSPCKNGATC-----SNSGQRSYT--CTC 62
QY 70 LPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATG 121
Db 63 RPYGTG-----VDCELELSE--CDSNFCRNGGCKDQEDGYXCLCPFG 103

RESULT 15
US-10-242-747-109
; Sequence 109, Application US/10242747
; Publication No. US20040005577A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207C1
; CURRENT APPLICATION NUMBER: US/10/242,747
```

```
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,881
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 109
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (187)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (188)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-242-747-109
```

```
Query Match 11.5%; Score 87; DB 15; Length 188;
Best Local Similarity 25.0%; Pred. No. 0.53;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

QY 14 WISAC-----AMLLCHGSLQHTFQOHLHRPEGGTCEVIAAHRCCNKNRIBERSQTVKCS 69
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
18 WQCTDEGWGGLFCDQDLNYC--THSPCKNGATC-----SNSGORSYT--CTC 62

QY 70 LRGKVAGTTNRPSVDASIVIGKWCMEPCLEGECKTLPDNSGWCATG 121
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63 RPYGTG-----VDCELELSE--CDSNPCRNGSGCKDQEDGYXCLCPPG 103
```

Search completed: April 8, 2004, 13:23:16
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 13:14:59 ; Search time 39 Seconds

(without alignments)
1075.998 Million cell updates/sec

Title: US-09-763-335-2

Perfect score: 754

Sequence: 1 MAMVSAMSWLYLWISACAM.....SGWMCATGNKIKTRIHPT 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTEMBL_25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriaph:*

17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	754	100.0	133	4	Q725A9
2	750	99.5	133	11	Q7TPG8
3	746	98.9	133	4	Q8TCL8
4	443	58.8	135	11	Q7TPG5
5	438	58.1	140	6	Q9NOD3
6	436.5	57.9	140	4	Q96LR4
7	435	57.7	140	6	Q95K92
8	432	57.3	131	4	Q8N3H0
9	429.5	57.0	132	11	Q7TPG6
10	429	56.9	131	11	Q7TPG7
11	415	55.0	126	11	Q8BV02
12	411.5	54.6	123	4	Q725A7
13	247.5	32.8	125	4	Q725A8
14	247.5	32.8	125	11	Q9IWE9
15	241	32.0	132	4	Q8IXR8
16	241	32.0	132	11	Q8CLV6

17	241	32.0	165	4	095902	095902 homo sapien
18	221.5	29.4	169	4	Q725A6	Q725A6 homo sapien
19	106	14.1	3871	5	Q20911	Q20911 caenorhabdi
20	95.5	12.7	2516	11	Q7TQ52	Q7TQ52 mus musculu
21	95.5	12.7	2526	11	Q7TQ51	Q7TQ51 mus musculu
22	95.5	12.7	2531	11	Q8K428	Q8K428 mus musculu
23	95.5	12.7	2531	11	Q7TQ50	Q7TQ50 mus musculu
24	94.5	12.5	2352	5	O61240	O61240 halocynthia
25	93	12.3	2531	5	O16004	O16004 lytechinus
26	92.5	12.3	2146	5	Q9VC97	Q9VC97 drosophila
27	89.5	11.9	2428	5	Q8I6X6	Q8I6X6 boophilus m
28	88.5	11.7	1212	13	O42347	O42347 gallus gall
29	88.5	11.7	2447	13	O13149	O13149 fugu rubrip
30	87	11.5	597	11	O35727	O35727 mus musculu
31	85	11.3	669	4	O75441	O75441 homo sapien
32	85	11.3	1511	4	O75412	O75412 homo sapien
33	85	11.3	1587	4	O00508	O00508 homo sapien
34	84.5	11.2	2528	13	O8AXP0	O8AXP0 cynops pyrr
35	84.5	11.2	2656	5	Q9GNU3	Q9GNU3 paracentrot
36	84	11.1	450	11	Q8KOH9	Q8KOH9 mus musculu
37	84	11.1	844	11	Q7TSG9	Q7TSG9 mus musculu
38	84	11.1	1405	11	Q8VHS2	Q8VHS2 mus musculu
39	83.5	11.1	729	13	Q7T3M4	Q7T3M4 brachydanio
40	83	11.0	305	6	Q9N028	Q9N028 macaca fasc
41	83	11.0	609	11	Q80YC5	Q80YC5 mus musculu
42	83	11.0	651	10	Q9FJE2	Q9FJE2 arabidopsis
43	83	11.0	674	5	Q8T4N9	Q8T4N9 strongyloce
44	83	11.0	1065	11	Q810H2	Q810H2 mus musculu
45	82.5	10.9	111	12	Q8JKE6	Q8JKE6 ectromelia

ALIGNMENTS

RESULT 1

Q725A9 PRELIMINARY; PRT; 133 AA.

AC Q725A9; (TREMBLrel. 25, Created)

DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE TAFAL.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;

RT "TAFAL: A Novel Secreted Family with Homology to CC-chemokines.";

RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY325114; AAP92406.1; -.

SQ SEQUENCE 133 AA; 14901 MW; C5DD1C7E5F997386 CRC64;

Query Match	100.0%;	Score 754;	DB 4;	Length 133;
Best Local Similarity	100.0%;	Pred. No. 2.6e-80;		
Matches 133;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAMVSAMSWLYLWISACAMLLCHGSLQHTFQOHLHRLPEGGTCEVIAAHRCCNKNRIEE	60	
Db	1	MAMVSAMSWLYLWISACAMLLCHGSLQHTFQOHLHRLPEGGTCEVIAAHRCCNKNRIEE	60	
QY	61	RSQTVKCSCLPGKAVGATTRNRPSVCVDASIVIGKWCEMEPCLEGBECKTLPDSNGWMCAT	120	
Db	61	RSQTVKCSCLPGKAVGATTRNRPSVCVDASIVIGKWCEMEPCLEGBECKTLPDSNGWMCAT	120	
QY	121	GNKIKTRIHPT 133		
Db	121	GNKIKTRIHPT 133		
RESULT 2				
Q07TPG8				

RESULT 2
Q7TPG8

```

ID Q7PG8 PRELIMINARY; PRT; 133 AA.
AC Q7PG8;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE TAF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT "TAPA: A Novel Secreted Family with Homology to CC-chemokines.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY325120; AAF92412.1; -.
SQ SEQUENCE 133 AA; 14873 MW; C5C20C764F8E7386 CRC64;

Query Match 99.5%; Score 750; DB 11; Length 133;
Best Local Similarity 99.2%; Pred. No. 7.7e-80;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAMVMSWVLYLWISACAMLLCHGSLQHTFQQHHLRPEGTCVIAAHRCCNKNRIE 60
Db 1 MAMVMSWVLYLWISACAMLLCHGSLQHTFQQHHLRPEGTCVIAAHRCCNKNRIE 60
QY 61 RSQTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNGMWCAT 120
Db 61 RSQTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNGMWCAT 120
QY 121 GNKIKTRIHPT 133
Db 121 GNKIKTRIHPT 133

RESULT 3
Q8TCL8 PRELIMINARY; PRT; 133 AA.
AC Q8TCL8;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Hypothetical protein.
GN DXFP566B064.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Bloeker H., Boescher M., Brandt P., Mewes H.W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL713702; CAD28501.1; -.
KW Hypothetical protein.
SQ SEQUENCE 133 AA; 15030 MW; C5C8082F0F997386 CRC64;

Query Match 98.9%; Score 746; DB 4; Length 133;
Best Local Similarity 99.2%; Pred. No. 2.3e-79;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAMVMSWVLYLWISACAMLLCHGSLQHTFQQHHLRPEGTCVIAAHRCCNKNRIE 60
Db 1 MAMVMSWVLYLWISACAMLLCHGSLQHTFQQHHLRPEGTCVIAAHRCCNKNRIE 60
QY 61 RSQTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNGMWCAT 120
Db 61 RSQTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNGMWCAT 120
QY 121 GNKIKTRIHPT 133
Db 121 GNKIKTRIHPT 133

```

```

Db 121 GNKIKTRIHPT 133

RESULT 4
Q7TPG5 PRELIMINARY; PRT; 135 AA.
ID Q7TPG5;
AC Q7TPG5;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE TAF4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT "TAPA: A Novel Secreted Family with Homology to CC-chemokines.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY325123; AAF92415.1; -.
SQ SEQUENCE 135 AA; 15018 MW; 38C83231992D3907 CRC64;

Query Match 58.8%; Score 443; DB 11; Length 135;
Best Local Similarity 62.6%; Pred. No. 6.7e-44;
Matches 77; Conservative 17; Mismatches 25; Indels 4; Gaps 2;

QY 9 WVLYLWISACAMLLCHGSLQHT--FQHHHLRPEGTCVIAAHRCCNKNRIERSQTVK 66
Db 13 WVLYLWVCCVCLMGASSQHLRFGHAGHLIKP--GTCEVAVHRCNKNRIERSQTVK 70
QY 67 CSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNGMWCATGNKI 126
Db 71 CSCFPQVAGTTRAQSCVEAIVIEKWWCHMNPCLGEGDCKVLPDSSGWSGSSGNK 130
QY 127 TRI 129
Db 131 TKV 133

RESULT 5
Q9NOD3 PRELIMINARY; PRT; 140 AA.
ID Q9NOD3;
AC Q9NOD3;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
DE Unnamed protein product.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB045997; BAB01579.1; -.
SQ SEQUENCE 140 AA; 15640 MW; 173DA66AB6C97F03 CRC64;

Query Match 58.1%; Score 438; DB 6; Length 140;
Best Local Similarity 62.8%; Pred. No. 2.7e-43;
Matches 76; Conservative 14; Mismatches 31; Indels 0; Gaps 0;

QY 9 WVLYLWISACAMLLCHGSLQHTFQQHHLRPEGTCVIAAHRCCNKNRIERSQTVK 68
Db 18 WVFLAVLVVCCVCLMGASSQHLRFGHAGHGHQIQGTCEVAVHRCNKNRIERSQTVK 77
QY 69 CLPGKVGAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNGMWCATGNKI 128

```

```

Db      78 CFFGQVAGTTAQPSCVEASIVIQKWCNNPCLEGECKVLDPDYSWSSSGNKVKYTK 137
QY      129 I 129
Db      138 V 138

RESULT 6
Q96LR4 ID Q96LR4 PRELIMINARY; PRT; 140 AA.
AC Q96LR4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ25161 (Similar to expressed sequence AW049604)
DE (TAF4).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Suzuki Y., Hata H.,
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
RA Irie R., Otsubi T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,
RA Nagai K., Isogai T., Sugano S.;
RA "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RL "TAPA: A Novel Secreted Family with Homology to CC-chemokines.";
RT Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK057890; BAB71606.1; -
DR EMBL; BC031566; AAH31566.1; -
DR EMBL; AY325117; AAP92409.1; -
KW Hypothetical protein.
SQ SEQUENCE 140 AA; 15682 MW; 173254FB8B526009 CRC64;

Query Match 57.9%; Score 436.5; DB 4; Length 140;
Best Local Similarity 61.1%; Pred. No. 4e-43;
Matches 80; Conservative 16; Mismatches 30; Indels 5; Gaps 2;

QY 4 VSAMSVLYLWISACAMLLCHGSIQHTFQOH---HL--HRPEGTCEVIAHRCCKNRI 58
Db 8 VCAKSVLLSHWFLAYVLMVCCKLMSASSQHLRGHAGHQIKQGTCEVAVHRCCKNRI 67

QY 59 EERSQIVKSCFLPGKVGAGTTNRPPSCVDASIVIGKWCNMEPCLEGECKTLDPDYSWSS 118
Db 68 EERSQIVKSCFFGQVAGTTAQPSCVEASIVIQKWCNNPCLEGECKVLDPDYSWSS 127

QY 119 ATGNKIKTKRI 129
Db 128 SSGNKVKTKV 138

RESULT 7
Q95K92 ID Q95K92 PRELIMINARY; PRT; 140 AA.
AC Q95K92;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Medulla oblongata;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB063062; BAB60784.1; -
KW Hypothetical protein.
SQ SEQUENCE 140 AA; 15654 MW; 17324670B6C97F19 CRC64;

Query Match 57.7%; Score 435; DB 6; Length 140;
Best Local Similarity 62.0%; Pred. No. 6.1e-43;
Matches 75; Conservative 15; Mismatches 31; Indels 0; Gaps 0;

QY 9 WLYLWISACAMLLCHGSIQHTFQOHHLHRPEGTCEVIAHRCCKNRIEERSQIVKCS 68
Db 18 WLFVLYVLMVCCKLMSASSQHLRGHAGHQIKQGTCEVAVHRCCKNRIEERSQIVKCS 77

QY 69 CLPGKVAGTTNRPPSCVDASIVIGKWCNMEPCLEGECKTLDPDYSWSSCATGNKIKTKR 128
Db 78 CFFGQVAGTTAQPSCVEASIVIQKWCNNPCLEGECKVLDPDYSWSSSGNKVKYTK 137

QY 129 I 129
Db 138 V 138

RESULT 8
Q8N3H0 ID Q8N3H0 PRELIMINARY; PRT; 131 AA.
AC Q8N3H0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (TAPA2).
GN DKFZF761E1217.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RA Ansoerke W., Witkner U., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RL "TAPA: A Novel Secreted Family with Homology to CC-chemokines.";
RT Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834160; CAD38865.1; -
DR EMBL; AY325115; AAP92407.1; -
KW Hypothetical protein.
SQ SEQUENCE 131 AA; 14620 MW; 6D7E0DB2E59965E4 CRC64;

Query Match 57.3%; Score 432; DB 4; Length 131;
Best Local Similarity 61.2%; Pred. No. 1.3e-42;
Matches 74; Conservative 14; Mismatches 23; Indels 10; Gaps 1;

QY 9 WLYLWISACAMLLCHGSIQHTFQOHHLHRPEGTCEVIAHRCCKNRIEERSQIVKCS 68
Db 19 FIVTLW-----GKVSSANHHKAAHVKTGTCEVVALHRCCKNRIEERSQIVKCS 68

QY 69 CLPGKVAGTTNRPPSCVDASIVIGKWCNMEPCLEGECKTLDPDYSWSSCATGNKIKTKR 128
Db 69 CFFGQVAGTTAQPSCVDASIVIQKWCNNPCLEGECKVLDPDYSWSSSGNKVKYTKR 128

```

```
QY 129 I 129
Db 129 V 129

RESULT 9
Q7TPG6 PRELIMINARY; PRT; 132 AA.
AC Q7TPG6;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE TAPAF3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT "TAPA: A Novel Secreted Family with Homology to CC-chemokines.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY325122; AAP92414.1; -.
SQ SEQUENCE 132 AA; 14426 MW; 2D60C45AE1BA3F00 CRC64;

Query Match 57.0%; Score 429.5; DB 11; Length 132;
Best Local Similarity 61.5%; Pred. No. 2.5e-42;
Matches 83; Conservative 13; Mismatches 16; Indels 23; Gaps 4;

QY 5 SAMSWSLYL---WISACAMLLCHGSLQHTFQOHLHRP-----EGGTCVIAAHRCCN 54
Db 9 SAGSWVLALCLAWLWCP-----ASASLQ-----PPTSAVLVKGTCEVIAAHRCCN 55

QY 55 KNRIERSQTVKSCULPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNS 114
Db 56 RNRIERSQTVKSCULPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNS 115

QY 115 GWMCATGNKIKTKTRI 129
Db 116 GWSGSGHKVKTKTKV 130

RESULT 10
Q7TPG7 PRELIMINARY; PRT; 131 AA.
AC Q7TPG7;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE TAPAF2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT "TAPA: A Novel Secreted Family with Homology to CC-chemokines.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY325121; AAP92413.1; -.
SQ SEQUENCE 131 AA; 14647 MW; FE70D61916A95F47 CRC64;

Query Match 56.9%; Score 429; DB 11; Length 131;
Best Local Similarity 61.2%; Pred. No. 2.8e-42;
Matches 74; Conservative 12; Mismatches 25; Indels 10; Gaps 1;

QY 9 WVLWLSACAMLLCHGSLQHTFQOHLHRPEGGTCVIAAHRCCNRIEERSQTVKCS 68
Db 19 FIVTLW-----GKAVSSANHHKAHVHTGTCVIAAHRCCNKNRIEERSQTVKCS 69

QY 69 CLPGKVAGTTNRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATGNKIKTKR 128

Db 69 CFPQGVAGTTNRAAPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWSGSGNKKTKR 128

QY 129 I 129
Db 129 V 129

RESULT 11
Q8BV02 PRELIMINARY; PRT; 126 AA.
AC Q8BV02;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE CDNA FLJ25161 FIS.
GN C130034I18RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK081531; BAC38247.1; -.
DR MGD; MGI:2444563; C130034I18RIK.
SQ SEQUENCE 126 AA; 13975 MW; D2AA1F362D44ACBA CRC64;

Query Match 55.0%; Score 415; DB 11; Length 126;
Best Local Similarity 62.9%; Pred. No. 1.2e-40;
Matches 73; Conservative 14; Mismatches 25; Indels 4; Gaps 2;

QY 9 WVLWLSACAMLLCHGSLQHT---FQOHLHRPEGGTCVIAAHRCCNRIEERSQTVK 66
Db 13 WLVLTYVLVVCCKLMSASSQHLRGHAGHLLKP--GTCEVAVHRCNKNRIEERSQTVK 70

QY 67 CSCLPGKVAGTTNRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATGN 122
Db 71 CSCFFQGVAGTTNRAAPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWSGSGN 126

RESULT 12
Q7Z5A8 PRELIMINARY; PRT; 133 AA.
AC Q7Z5A8;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE TAPAF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT "TAPA: A Novel Secreted Family with Homology to CC-chemokines.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY325116; AAP92408.1; -.
SQ SEQUENCE 133 AA; 14776 MW; C074C703817D008D CRC64;

Query Match 54.6%; Score 411.5; DB 4; Length 133;
Best Local Similarity 58.5%; Pred. No. 3.2e-40;
Matches 76; Conservative 15; Mismatches 26; Indels 13; Gaps 2;

QY 5 SAMSWSYL-----YLWTSACAMLLCHGSLQHTFQOHLHRPEGGTCVIAAHRCCNRIE 59
Db 5 SAMSWSYL-----YLWTSACAMLLCHGSLQHTFQOHLHRPEGGTCVIAAHRCCNRIE 59
```

Db 10 STGWLALCLANLWHLTLAALQPPTATVLVQ-----GTCEVIAHRCNKNRIE 61
QY 60 ERSQTVKSCULPGKVGAGTRNRPSVDASIVIGKWCMEPCLEGECKTLDPNSGMCA 119
Db 62 ERSQTVKSCULPGKVGAGTRNRPSVDASIVIGKWCMEPCLEGECKTLDPNSGMCA 121
QY 120 TGNKIKITRI 129
Db 122 SGHKVKTIV 131

RESULT 13
Q7Z5A7 PRELIMINARY; PRT; 125 AA.
AC Q7Z5A7
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE TAFAS.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tang Y. T., Entage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT "TAFAS: A Novel Secreted Family with Homology to CC-chemokines";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY325118; AAP92410.1; -;
SQ SEQUENCE 125 AA; 13592 MW; A8BCA8D09F86404A CRC64;

Query Match 32.8%; Score 247.5; DB 4; Length 125;
Best Local Similarity 44.3%; Pred. No. 4.7e-21;
Matches 54; Conservative 14; Mismatches 45; Indels 9; Gaps 3;
QY 13 LWSACAMLLCHGSLQHTFOQHLHRRPEG-----GTCEVIAHRCNKNRIEERSQTVKCS 68
Db 7 LWSACAMLLCHGSLQHTFOQHLHRRPEG-----GTCEVIAHRCNKNRIEERSQTVKCS 62
QY 69 CLPGKVGAGTRNRPSVDASIVIGKWCMEPCLEGECKTLDPNSGMCA-TGNKIKIT 127
Db 63 CRKQIAGTTRARPACVDARIITKQWCDMLPCLEGGCDLLINRSQWTCTQPGRIKT 122
QY 128 RI 129
Db 123 TV 124

RESULT 14
Q91WE9 PRELIMINARY; PRT; 125 AA.
AC Q91WE9
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (TAFAS).
GN AW049604
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Tang Y. T., Entage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT "TAFAS: A Novel Secreted Family with Homology to CC-chemokines";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015306; AAH15306.1; -;

DR EMBL; AY325124; AAP92416.1; -;
DR MGD; MGI:2146182; AW049604.
KW Hypothetical protein.
SQ SEQUENCE 125 AA; 13592 MW; A8BCA8D09F86404A CRC64;

Query Match 32.8%; Score 247.5; DB 11; Length 125;
Best Local Similarity 44.3%; Pred. No. 4.7e-21;
Matches 54; Conservative 14; Mismatches 45; Indels 9; Gaps 3;
QY 13 LWSACAMLLCHGSLQHTFOQHLHRRPEG-----GTCEVIAHRCNKNRIEERSQTVKCS 68
Db 7 LWSACAMLLCHGSLQHTFOQHLHRRPEG-----GTCEVIAHRCNKNRIEERSQTVKCS 62
QY 69 CLPGKVGAGTRNRPSVDASIVIGKWCMEPCLEGECKTLDPNSGMCA-TGNKIKIT 127
Db 63 CRKQIAGTTRARPACVDARIITKQWCDMLPCLEGGCDLLINRSQWTCTQPGRIKT 122
QY 128 RI 129
Db 123 TV 124

RESULT 15
Q81XR8 PRELIMINARY; PRT; 132 AA.
AC Q81XR8
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to expressed sequence AW049604.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC039396; AAH39396.1; -;
SQ SEQUENCE 132 AA; 14300 MW; 16BFE9F574FBB066 CRC64;

Query Match 32.0%; Score 241; DB 4; Length 132;
Best Local Similarity 38.4%; Pred. No. 2.9e-20;
Matches 48; Conservative 23; Mismatches 42; Indels 12; Gaps 2;
QY 6 AMSWLVLWISACAMLLCHGSLQHTFOQHLHRRPEGSTCEVIAHRCNKNRIEERSQTV 65
Db 18 SMSSTFWAFWILASLIAYCS-----QLAAGTCEIVTLDRDSSQPRRTIARQTA 66
QY 66 KCSCLPGKVGAGTRNRPSVDASIVIGKWCMEPCLEGECKTLDPNSGMCA-TGNKI 124
Db 67 RCACRKGQIAGTTRARPACVDARIITKQWCDMLPCLEGGCDLLINRSQWTCTQPGRI 126
QY 125 KITRI 129
Db 127 KITTV 131

Search completed: April 8, 2004, 13:17:32
Job time : 40 secs